
CMS Internal Note

The content of this note is intended for CMS internal use and distribution only

15 September 2007

Jet and MET Performance in CMSSW_1_2_0

M. Vázquez Acosta

CERN, Geneva, Switzerland

L. Apanasevich, C. Dragoiu, A. Smoron, N. Varelas

University of Illinois at Chicago, Chicago, IL , USA

A. Bhatti

Rockefeller University, New York, NY, USA

R. Cavanaugh, B. Scurlock

University of Florida, Gainesville, FL , USA

R. Demina, J. Widawsky, M. Zieliński

University of Rochester, Rochester, NY, USA

G. Dissertori

ETH Zurich, Switzerland

R. Harris

Fermilab, Batavia, IL, USA

M. Jha

University of Delhi, Delhi, India

S. Petrushanko, O. Kodolova

Moscow State University, Moscow, Russia

F. Ratnikov

University of Maryland, College Park, MD, USA

Abstract

This note summarizes results of jet and Missing ET (MET) performance studies with CMSSW_1_2_0, completed for the Software and Detector Performance Validation (SDPV) exercise in March 2007.

The scope of the paper includes a brief overview of jet reconstruction in CMSSW_1_2_0; the derivation of MCJet energy corrections; a discussion of the impact of pileup on jet reconstruction, studies of jet energy response, energy and position resolutions, and jet efficiencies for various pileup samples; dijet balance approach to jet response equalization across the detector; and Missing ET reconstruction and performance.

1 Introduction

Majority of the results presented in this note have been obtained during the course of the Software and Detector Performance Validation (SDPV) effort. The main goal of the SDPV was to establish the then-current level of performance of CMSSW simulation and reconstruction and compare to selected benchmark results based on the ORCA software and published in the CMS Physics TDR [1]. Detailed comparisons between CMSSW and ORCA results can be found in the SDPV meeting presentations [2] and the final SDPV report to the collaboration [3].

In this note we report the jet and MET performance studies based on CMSSW_1_2_0 in order to establish a baseline for comparing results from future releases of CMSSW. However, for many plots we retain the original format intended for comparisons with ORCA and include selected direct comparisons or comments, when of particular interest. When making any detailed comparisons between CMSSW and ORCA one needs to keep in mind various differences in the respective choices of settings, including:

- Tune DWT for the Underlying Event parameters in PYTHIA has been used in CMSSW (older settings in ORCA)
- Different versions of GEANT, detector material description and geometry
- Differences in calorimeter simulation/digitization/calibration, different Shower Library for the Forward Hadron calorimeter (HF)
- In CMSSW, the HCAL energy is calculated from the sum of four time-slices (while two time-slices were used in ORCA)
- Pileup settings, as discussed below (ORCA studies typically corresponded to low-luminosity pileup, $2 \times 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$)
- Recombination E -scheme for jets is the default in CMSSW (E_T -scheme used for many ORCA studies)
- Scheme-B calorimeter cell thresholds for jet and MET reconstruction (and 0.5 GeV tower E_T cut for jets) are used in CMSSW. In ORCA, a zero suppression threshold of 0.5 GeV was applied on HCAL cell energies used in jet and MET reconstruction and additional cuts on tower $E > 0.8 \text{ GeV}$ and $E_T > 0.5 \text{ GeV}$ were commonly used for jets.

After a brief overview of jet reconstruction in CMSSW_1_2_0 (Section 2) and Monte Carlo samples employed (Section 3), we discuss the derivation of MCJet jet energy corrections (Section 4), impact of pileup on jet reconstruction (Section 5), jet energy response, energy resolutions and position resolutions for various pileup samples (Section 6), jet efficiencies (Section 7), dijet balance approach to jet response equalization across the detector (Section 8). Missing ET reconstruction and performance are discussed in detail in Section 9. The paper concludes with Section 10.

2 Jet Reconstruction in CMSSW_1_2_0

CMSSW_1_2_0 contains three jet reconstruction algorithms, each of which is able to create three “flavors” of jet objects depending on the type (“flavor”) of input objects. Jet reconstruction is a two-step procedure. In the first step, an arbitrary input collection is treated as a set of Lorentz vectors. Every Lorentz vector satisfying energy and/or E_T requirements is used by the jet clustering algorithm. In the second step, after jet clustering is completed, flavor specific information is extracted from objects contributing to the jet, and corresponding flavor-specific values are associated with the jet.

Three jets flavors are available:

- *BasicJets* are produced from an arbitrary set of constituents. No specific constituent information is used, the produced jet contains only kinematics parameters and back references to the jet constituents.

- *CaloJets* are produced from constituents which are CaloTowers. Every CaloTower is built from one or more HCAL cells and corresponding ECAL crystals. In the barrel, a single projective Calorimeter Tower is formed from the unweighted sum of energy deposits in 5×5 ECAL crystals and in a single HCAL cell. In the forward regions, a more complex association of ECAL crystals with HCAL cells is required. In order to reduce the effects of electronic noise in CMSSW, so-called “Scheme-B” thresholds [4] were applied (by default) to all Calorimeter Towers:

Scheme	HB	HO	HE	Σ EB	Σ EE
A	0.70	0.85	0.90	0.20	0.45
B	0.90	1.10	1.40	0.20	0.45
C	1.20	1.30	1.80	0.20	0.45

Table 1: Threshold values (in GeV) for noise suppression schemes A, B and C. Columns labeled with Σ EB and Σ EE refer to thresholds on the sum of energy deposits in EB and EE crystals corresponding to a Calorimeter Tower.

In addition to basic kinematics parameters, CaloJets contain specific information collected from contributing CaloTowers, like energies deposited in different ECAL and HCAL detector regions, fractions of electromagnetic and hadronic energies of the jet, maximum electromagnetic and hadronic energies deposited in one tower etc.

- *GenJets* are produced from constituents which are Monte-Carlo generator level particles. Only particles marked as stable on the generator level contribute to a GenJet. A GenJet contains information about basic kinematic variables and energy contributions from electromagnetic particles, hadrons, and invisible particles contributing to the jet eg., neutrinos.

Three jet clustering algorithms are used in CMSSW_1_2_0 (see [1, 5] for additional details and a discussion of ORCA results):

- *Iterative Cone* algorithm ported from ORCA. A seed threshold is applied, input objects are assigned to a jet that is found first and then constituents are excluded from the input list; no merging/splitting is applied.
- *Midpoint Cone* algorithm has been developed for and is widely used in Tevatron Run II jet reconstruction. In CMSSW it is ported directly from the CDF code. Seed threshold is applied and proto-jet clusters are found without removing constituents that have been already used. A splitting-merging step is invoked for the produced proto-jet clusters, controlled by the Overlap Threshold parameter. The Search Cone step was not applied.
- k_T is an implementation of the standard algorithm ported using the original code [6]. The CMSSW_1_2_0 implementation of the k_T algorithm had known problems and hence it has been substituted by the external original k_T code in following releases and finally superseded by the FastJet implementation [7].

Table 2 summarizes the parameters used for CaloJet reconstruction in CMSSW_1_2_0.

	Iterative Cone	Midpoint Cone		k_T
Alias	IC5	MC5	MC7	k_T
Size	$R = 0.5$	$R = 0.5$	$R = 0.7$	$D = 1$
Tower thresholds	$E_T > 0.5 \text{ GeV}$	$E_T > 0.5 \text{ GeV}$		$E > 1 \text{ GeV}$
Cell thresholds	Scheme B			
Seed threshold	$E_T > 1 \text{ GeV}$	$E_T > 1 \text{ GeV}$		NA
Other parameters	NA	Overlap Threshold = 0.75 maxPairSize = 2 maxIterations = 100		NA

Table 2: Parameters for CaloJet reconstruction algorithms in CMSSW_1_2_0.

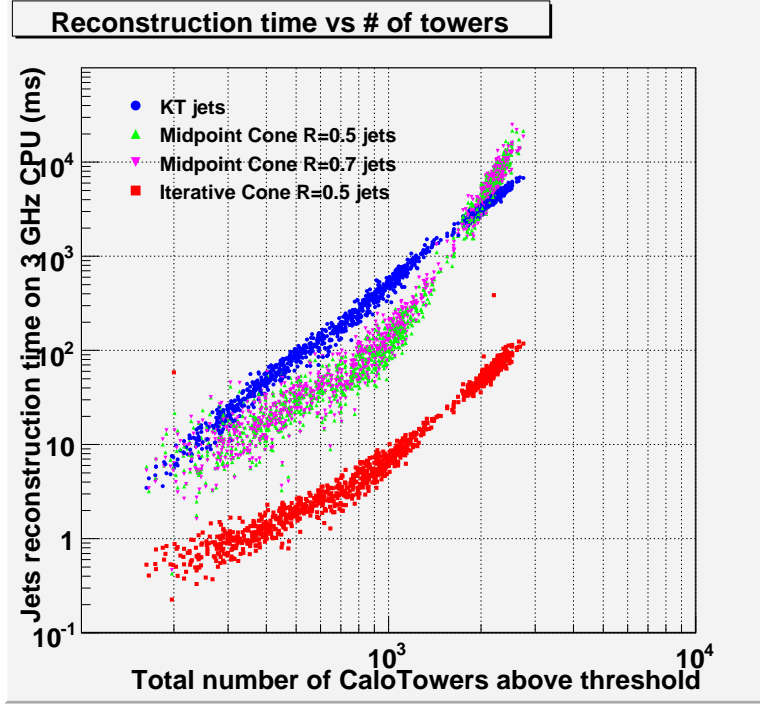


Figure 1: CPU time necessary for different algorithms to reconstruct CaloJets as a function of the total number of CaloTowers above threshold in the event. The range $N_{\text{CaloTower}} < 700$ corresponds to low luminosity, the range $500 < N_{\text{CaloTower}} < 1500$ corresponds to five extra minimum bias interactions per event, and the range $1500 < N_{\text{CaloTower}} < 3000$ corresponds to 25 extra minimum bias interactions per event.

The Iterative Cone algorithm is the fastest one. The k_T algorithm is well defined theoretically, but is also the slowest due to the N^3 dependency on the total number of input objects. However, the new FastJet implementation of k_T , available since CMSSW_1.3_x, is much faster, comparable in speed to the Iterative Cone algorithm. Figure 1 illustrates the CPU time necessary for different algorithms to reconstruct jets for different input multiplicities.

3 CMSSW_1.2.0 QCD Samples

This analysis employs Monte Carlo QCD event samples produced for the SDPV exercise using CMSSW_1.2.0, and a few special test samples as described below. The particle-level events were generated with PYTHIA 6.227 using the Tune DWT for Underlying Event parameters [8]. The CMS detector simulation as implemented in CMSSW_1.2.0 based on the GEANT4 package was used to simulate passage of particles through the detector and the energy deposits in the sensitive volumes. Unless stated otherwise, the results presented below have been derived from samples without pileup (PU). QCD dijet samples with no PU were generated in 21 bins of the momentum transfer in the parton hard-scatter, \hat{p}_T , which span the full kinematic range [9].

We include a brief discussion of various PU settings for CMSSW_1.2.0 and illustrate the impact of PU on some selected results. To simulate additional proton-proton interactions in a beam crossing from PU, the signal events were mixed with a random number of minimum bias events in one crossing. The minimum bias events were generated with PYTHIA as inclusive QCD events. The Poisson distribution with an average of five was used to simulate PU, corresponding to the so-called “low-lumi” conditions. The QCD samples with in-time PU were part of the official OnSel production, and were available for $\hat{p}_T < 800$ GeV (samples with no PU were used for higher \hat{p}_T values to extend the range of coverage in some plots, as needed).

Finally we also used small samples with full (in-time plus out-of-time) low-lumi PU from a private production for $50 < \hat{p}_T < 80$ GeV [10], and with a HF Shower Library fix corresponding to CMSSW_1.2.2 (for $80 < \hat{p}_T < 120$ GeV and no PU).

4 MC Jet Corrections

Here we describe the MC Jet corrections which correct the calorimeter level jet (CaloJet) to have the same E_T as the particle level jet (GenJet). These corrections are based on a QCD dijet Monte Carlo simulation using PYTHIA and the CMS detector simulation with no PU. The methodology was developed previously for OSCAR/ORCA [5] and ported to CMSSW.

4.1 Jet Response

Measurements of the jet response were made with the module SimJetResponseAnalysis in the MCJet package of the JetMETCorrections subsystem of CMSSW. Jets were reconstructed using the iterative cone and midpoint cone algorithms in CMSSW. We consider all GenJets in the event and match each GenJet with the closest CaloJet which minimizes

$$\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}, \quad (1)$$

where η and ϕ denote jet pseudorapidity and azimuthal angle, respectively. If the closest CaloJet is not within $\Delta R = 0.25$, the GenJet is discarded. For all passing GenJets we measure

$$\text{Jet Response} = \frac{\text{CaloJet } E_T}{\text{GenJet } E_T} \quad (2)$$

in the bins of GenJet E_T and CaloJet $|\eta|$ listed in table 3 and table 4:

10	12	15	20	27	35	45	57	72	90	120
150	200	300	400	550	750	1000	1400	2000	2900	4500

Table 3: Bin edges of GenJet E_T in GeV used for measurements of the jet response.

0.0	0.226	0.441	0.751	0.991	1.260	1.496	1.757	2.046
2.295	2.487	2.690	2.916	3.284	4.0	4.4	4.8	

Table 4: Bin edges of CaloJet $|\eta|$ used for measurements of the jet response.

Example histograms of the jet response are shown in Fig 2. Notice that for each η region (row) the response increases and the resolution improves with increasing E_T (column). To determine the peak of the jet response, the most probable value, we have fit each of the histograms in Fig. 2 with Gaussians in the interval $\pm 1\sigma$ from the peak. If a full Gaussian fit is used instead of $\pm 1\sigma$, the mean value of the Gaussian increases by less than 3% depending on E_T . We use the mean value of the $\pm 1\sigma$ Gaussian fits to define the peak jet response at the average GenJet E_T in a specific bin.

Example plots of peak jet response as a function of average GenJet E_T are shown in Fig. 3. The response increases smoothly with E_T . For each of the 16 bins of CaloJet $|\eta|$, we fit the response with the same parameterization used for ORCA [5]. The parameterization is compared to the response points in Fig. 3.

Changes in jet response with different versions of the CMS detector simulations are illustrated in Fig. 4. The CMSSW_1_2_0 response is compared with both CSA06 [11] and ORCA [5] samples. CSA06 used CMSSW_0_8_3 simulation and CMSSW_1_0_3 reconstruction. We see that the CMSSW_1_2_0 response is higher than CSA06 in the barrel and endcap. The following changes are known to have occurred and contribute to that difference

- The HCAL Endcap (HE) response in CSA06 was low by 35% because the reconstruction used the same calorimeter sampling fraction as the HCAL Barrel (HB), and this was fixed in time for the CMSSW_1_2_0 release.
- The RecHit energy in HB and HE was increased by 3% in CMSSW_1_2_0 to account for signal lost outside of the 4 time-slice integration window, and this correction was not present in CSA06.

The response in CMSSW_1_2_0 remains different than seen with ORCA by a few percent in the Barrel and Endcap, and by around 25% in the Forward due to the use of an uncalibrated shower library.

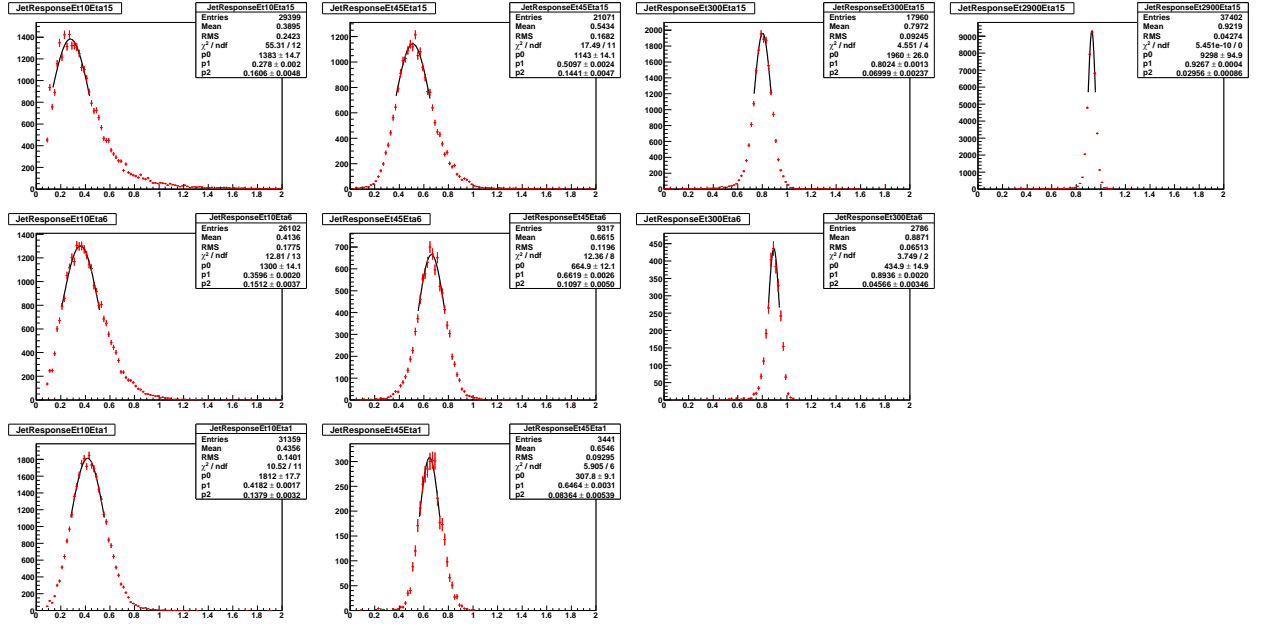


Figure 2: Jet response distributions and fits for the iterative cone algorithm with cone size $R = 0.5$. The number of jets as a function of jet response (points) is compared to a Gaussian fit (curve) in the interval $\pm 1\sigma$ from the peak response. The three rows of plots, from top to bottom, are for the following regions of CaloJet η : $0 < |\eta| < 0.226$, $2.295 < |\eta| < 2.487$, and $4 < |\eta| < 4.4$. The four columns of plots, from left to right, are for the following regions of GenJet E_T : $10 < E_T < 12$, $45 < E_T < 57$, $300 < E_T < 400$, and $2900 < E_T < 4500$ GeV.

4.2 MCJet Energy Corrections

The jet energy correction, k , is defined as

$$k = \frac{1}{\text{Jet Response}}. \quad (3)$$

It is a multiplicative correction: the CaloJet Lorentz vector p is multiplied by the jet correction to obtain a corrected CaloJet Lorentz vector p' :

$$p' = kp \quad (4)$$

The parameterized jet response as a function of GenJet E_T in 16 slices of CaloJet $|\eta|$ is used as input to the MC Jet package in the JetMETCorrections subsystem of CMSSW.

The software applies a simple iteration procedure to derive the jet correction as a function of the observed CaloJet E_T from the input Jet Response which is a function of true GenJet E_T . Let i be the iteration number, then k_i is the correction obtained in the i^{th} iteration, and is equal to

$$k_i = \frac{1}{\text{Jet Response}(\text{CaloJet } E_T \times k_{i-1})} \quad (5)$$

where $k_0 = 1$. In Eq. 5 we are substituting an approximation for the GenJet E_T into the Jet Response function of GenJet E_T , and with each iteration the approximation becomes more precise. The software iterates ten times to obtain a value of the jet correction as a function of CaloJet E_T . In other words, for each reconstructed CaloJet the software solves the non-linear equation

$$\frac{\text{CaloJet } E_T}{\text{GenJet } E_T} = \text{Jet Response}(\text{GenJet } E_T) \quad (6)$$

using a simple iteration procedure.

The jet correction as a function of CaloJet E_T is obtained for the bins of CaloJet $|\eta|$ given above. This means that for a fixed CaloJet E_T the correction is held constant within the bin of CaloJet $|\eta|$. The correction changes as a

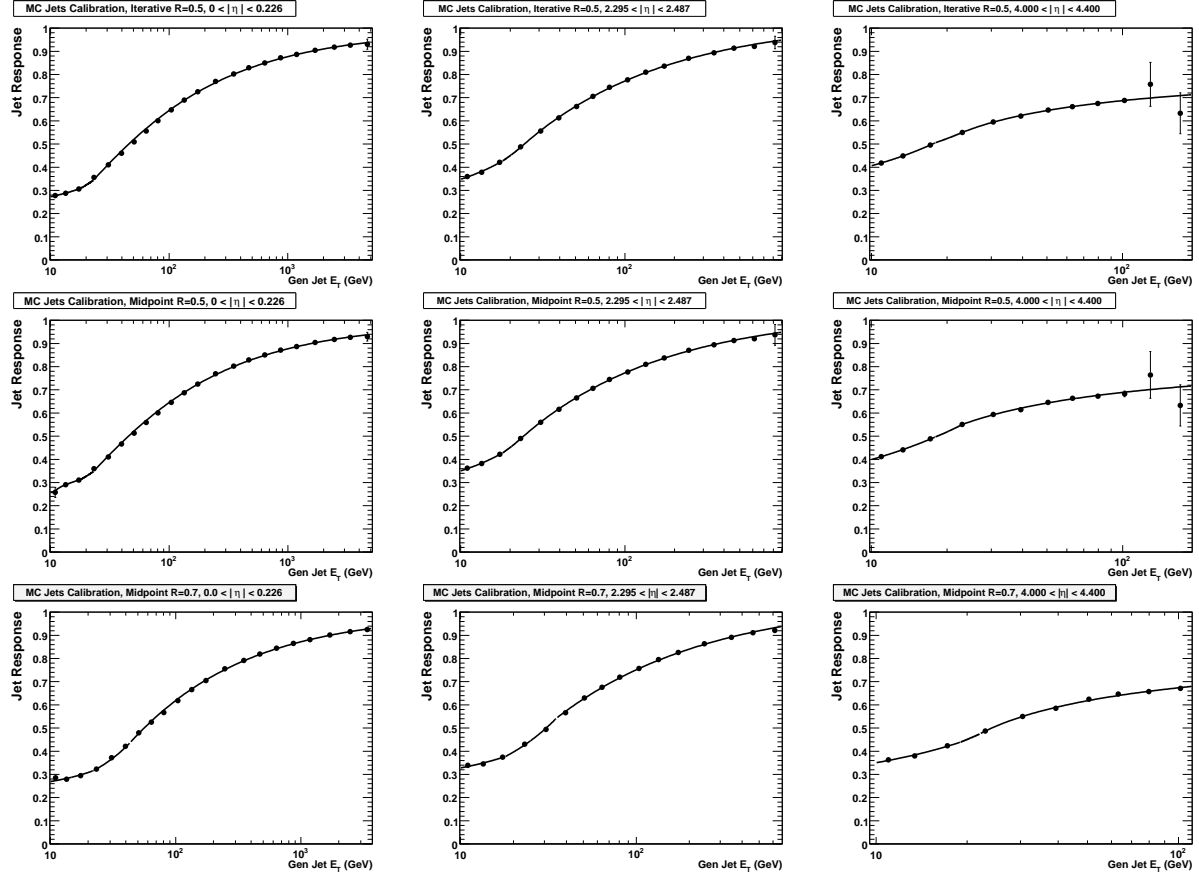


Figure 3: The jet response as a function of GenJet E_T (points) is compared to a parameterization (curve). The three rows of plots, from top to bottom, are for the iterative cone algorithm with a cone size $R = 0.5$, and for the midpoint cone algorithm with a cone size of $R = 0.5$ and $R = 0.7$. The three columns of the plot, from left to right, are for the following regions of CaloJet η : $0 < |\eta| < 0.226$, $2.295 < |\eta| < 2.487$, and $4 < |\eta| < 4.4$.

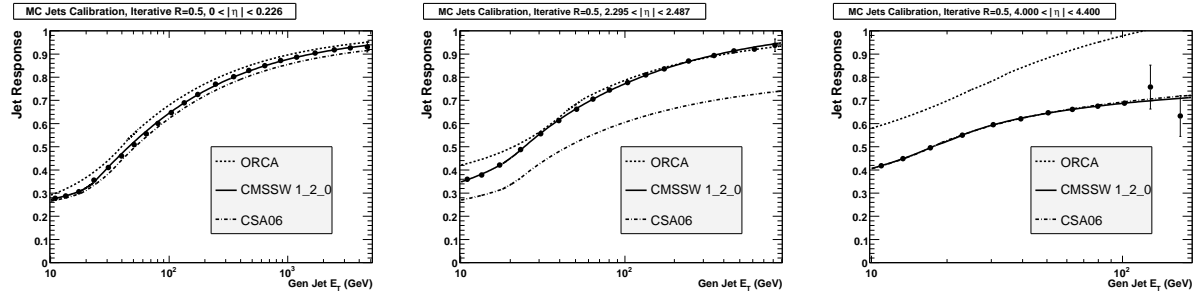


Figure 4: The jet response as a function of GenJet E_T from ORCA (dashed curve), CSA06 (dot-dashed curve), and CMSSW_1_2_0 (points and solid curve). The three plots, from left to right, are for the following regions of CaloJet η : $0 < |\eta| < 0.226$, $2.295 < |\eta| < 2.487$, and $4 < |\eta| < 4.4$.

function of $|\eta|$ in discrete jumps, as one moves from one bin of CaloJet $|\eta|$ to the next. We plan to replace this in the future with a correction that varies smoothly as a function of CaloJet η .

In Fig. 5 we show the jet correction as a function of CaloJet E_T in three bins of CaloJet $|\eta|$. Inversely to the behavior of the jet response, the jet correction decreases with increasing jet E_T , asymptotically approaching 1 in the barrel and the endcap at high CaloJet E_T .

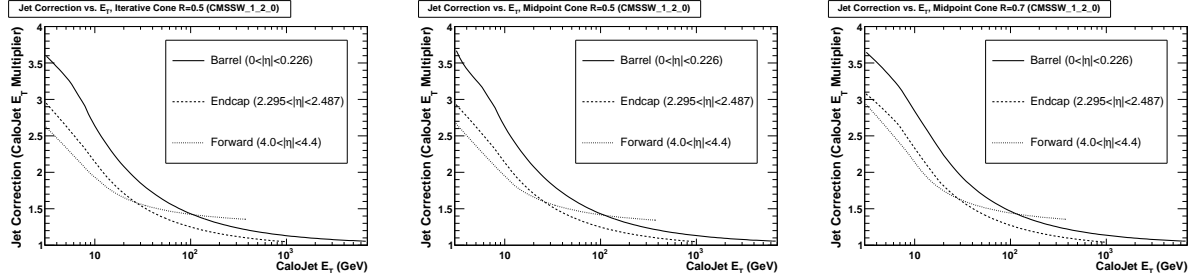


Figure 5: The jet correction as a function of CaloJet E_T for the iterative cone algorithm with a cone size of $R = 0.5$ (left plot), and the midpoint cone algorithm with a cone size of $R = 0.5$ (middle plot) and $R = 0.7$ (right plot).

4.3 Corrected Jet Response

As a closure test, the jet correction procedure described above was used in the MC Jet package to produce collections of corrected CaloJets. The corrected jet collections were then input to the same SimJetResponseAnalysis module used to measure the raw jet response. The corrected jet response

$$\text{Corrected Jet Response} = \frac{\text{Corrected CaloJet } E_T}{\text{GenJet } E_T} \quad (7)$$

was histogrammed in the same bins of GenJet E_T and CaloJet $|\eta|$.

Example histograms of corrected jet response are shown in Fig. 6. The corrected jet response peaks near 1, particularly at high GenJet E_T , indicating that the correction is working. At very low values of GenJet E_T , typically less than 30 GeV, the jet resolution is very poor, and the correction made to the jet varies significantly. As a result the shape of the corrected jet response at low GenJet E_T in Fig. 6 is slightly different from the shape of the uncorrected jet response in Fig. 2. Consequently, for these very low energy jets, the correction does not work as well. To quantify how well the correction is working, we have determined the peak of the corrected jet response by fitting each histogram with Gaussians in the interval $\pm 1\sigma$ from the peak, as shown in the examples of Fig. 6.

Example plots of the corrected jet response as a function of average GenJet E_T are shown in Fig. 7. The correction is working to within $\approx 1\%$ accuracy for GenJet $E_T > 30$ GeV. The accuracy is significantly worse (the corrected jet response peaks around 2% to 10% above the expected value of 1) at lower E_T , particularly in the barrel where the jet resolution is worse than in the endcap or forward regions for fixed E_T .

An example of the jet response as a function of η before and after jet corrections is shown in Fig. 8. Before jet corrections the plot shows the response variations of the CMS detector simulation as a function of η . After the jet corrections are applied, the response is reasonably flat around 1. The vertical dotted lines show the edges of the $|\eta|$ bins in Table 4, while the points are the jet response measured in bins equal to the CaloTower η segmentation.

For completeness we note that Fig. 8 was made in bins of GenJet p_T (not E_T) and in this figure we are plotting the mean of the jet response distribution, not the peak of a $\pm 1\sigma$ Gaussian fit. Also note that the distribution is asymmetric in η for $|\eta| > 3$ due to a known problem with the η values returned by the HF simulation; this was fixed in CMSSW_1_2_2 and higher versions.

5 Pileup Considerations for Jet Reconstruction

At the design luminosity ($10^{34} \text{ cm}^{-2}\text{s}^{-1}$), the LHC is expected to deliver on average about 20 proton-proton interactions per beam crossing. Each event will receive contributions both from particles produced in the trigger

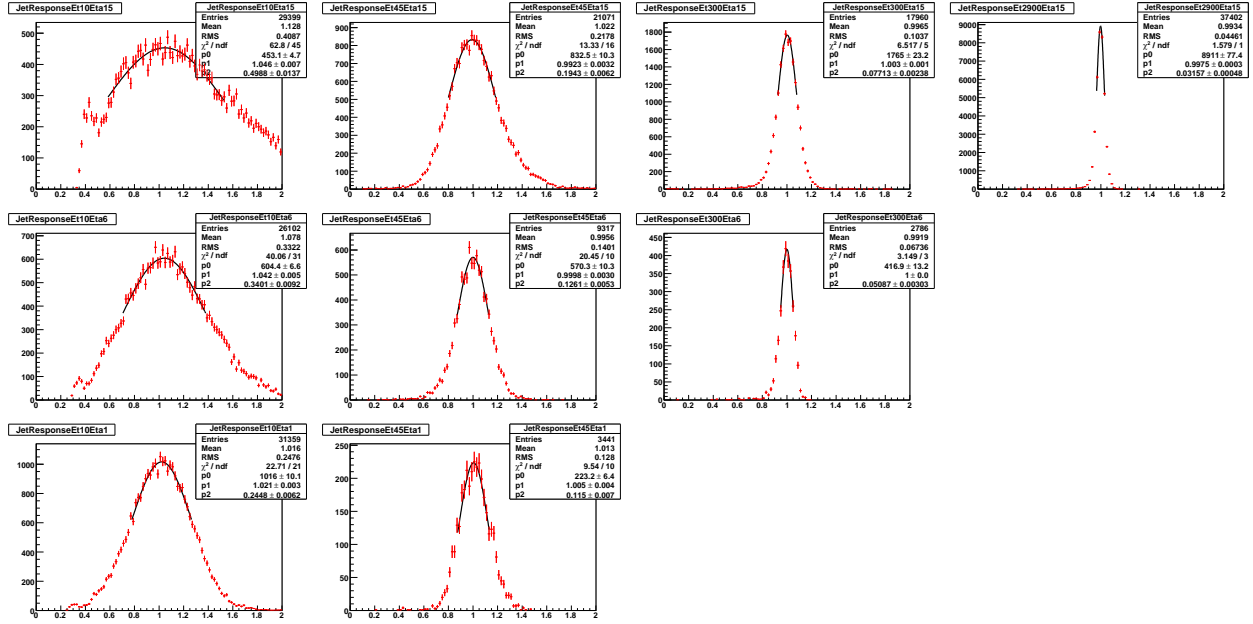


Figure 6: Corrected jet response distributions and fits for the iterative cone algorithm with cone size $R = 0.5$. The number of jets as a function of corrected jet response (points) is compared to a Gaussian fit (curve) in the interval $\pm 1\sigma$ from the peak response. The three rows of plots, from top to bottom, are for the following regions of CaloJet η : $0 < |\eta| < 0.226$, $2.295 < |\eta| < 2.487$, and $4 < |\eta| < 4.4$. The four columns of plots, from left to right, are for the following regions of GenJet E_T : $10 < E_T < 12$, $45 < E_T < 57$, $300 < E_T < 400$, and $2900 < E_T < 4500$ GeV.

crossing (in-time PU) and from particles produced in the adjacent crossings (out-of-time PU). Pileup of particles from different interactions will produce energy clusters in the calorimeter which will affect jet reconstruction. Below, we demonstrate the effect of in-time and full (in-time + out-of-time) PU on jet pseudorapidity distributions.

Since PU events contribute mostly low energy particles, one of the most striking effects of PU is on the total E_T observed in the calorimeter; for the $50 < \hat{p}_T < 80$ GeV samples, the average total E_T is ≈ 180 GeV for the no-PU case, 290 GeV for in-time PU and 410 GeV for full PU. Consequently, the impact of PU on jet reconstruction is much larger at low jet E_T than at high E_T (the effect on MET reconstruction is discussed in detail in Section 9).

The variation of the pseudorapidity distribution of CaloJets with uncorrected $10 < p_T < 20$ GeV in the QCD sample with $50 < \hat{p}_T < 80$ GeV for no PU, in-time PU and full PU is illustrated in Fig. 9. The overall low number of CaloJets being reconstructed in the calorimeter, as compared to the generated jet distributions, is due to the low average calorimeter response in this range of jet p_T . The forward calorimeter doesn't suffer from out-of-time PU due to its very short response time. This can be seen in the Fig. 9(b) and Fig. 9(c). The endcap towers near the HE/HF boundary have the largest transverse size in the calorimeter and tend to accumulate significant energy flow from PU events. It is due to this effect that the “horns” are visible in the full PU sample (Fig. 9(c)) around $|\eta| = 3$.

Figure 10 shows the η distribution for jets with $40 < p_T < 60$ GeV in the $50 < \hat{p}_T < 80$ GeV samples. It is evident from the Figs. 9(c) and 10(c) that the impact of PU on η distributions of Calo jets is reduced at higher p_T .

6 Jet Response and Resolutions

The goal of the following jet response study has been to provide plots in the format corresponding to previous results from ORCA; these plots were a part of the SDPV effort. Clearly, there is a significant overlap with the results and methods presented in Section 4. Below, we also include results for various PU conditions using the same analysis code and plot formats.

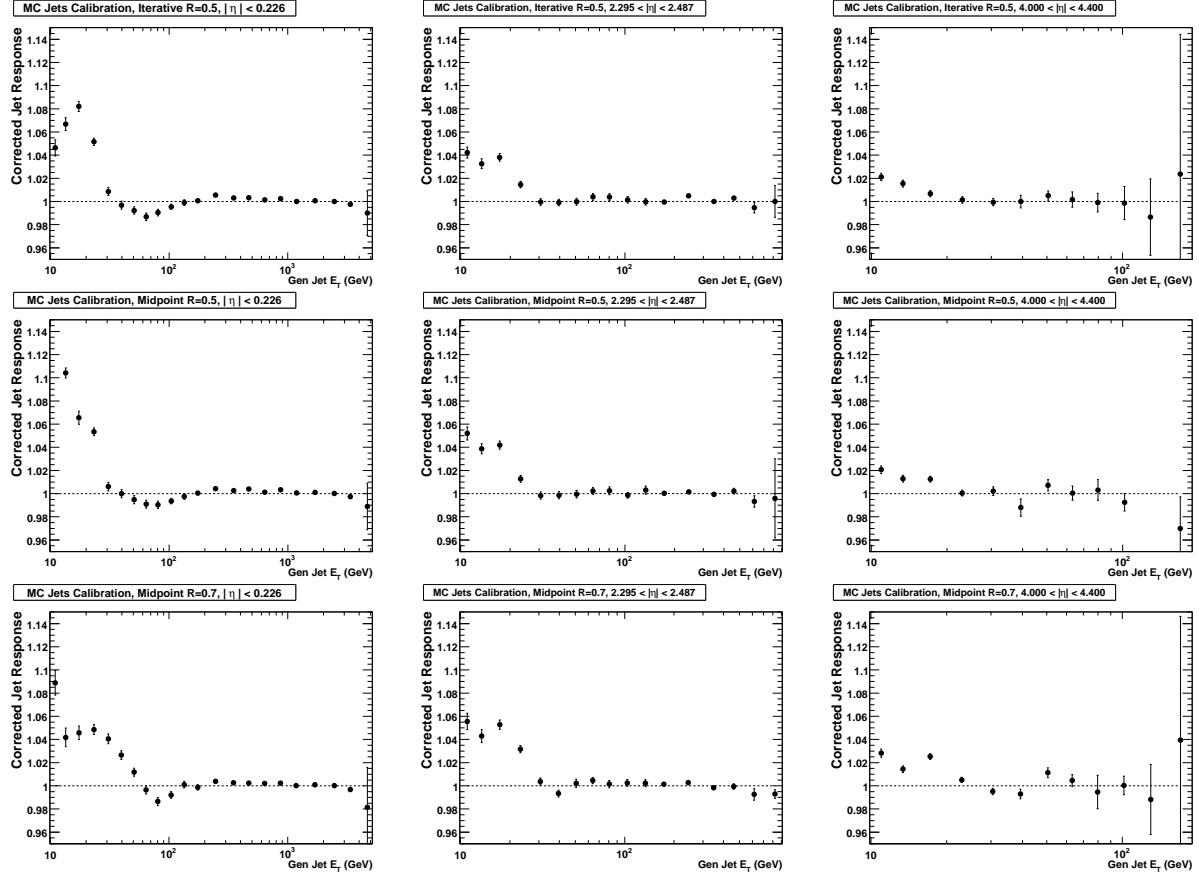


Figure 7: The corrected jet response as a function of GenJet E_T (points). The three rows of plots, from top to bottom, are for the iterative cone algorithm with a cone size $R = 0.5$, and for the midpoint cone algorithm with a cone size of $R = 0.5$ and $R = 0.7$. The three columns of plot, from left to right, are for the following regions of CaloJet η : $0 < |\eta| < 0.226$, $2.295 < |\eta| < 2.487$, and $4 < |\eta| < 4.4$.

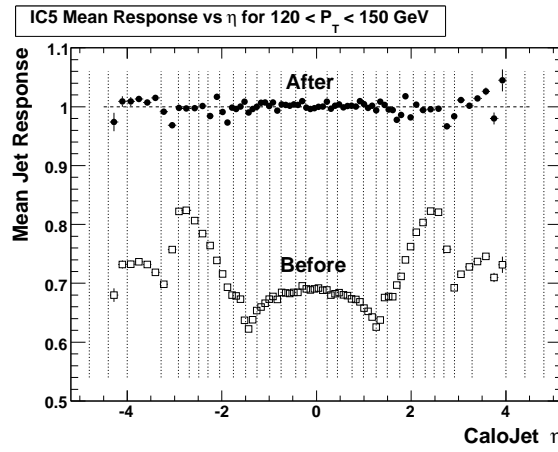


Figure 8: The jet response as a function of CaloJet η before jet corrections (boxes) and after jet corrections (closed circles), for GenJet p_T in the range $120 < p_T < 150$ GeV for the Iterative cone algorithm with cone size $R = 0.5$.

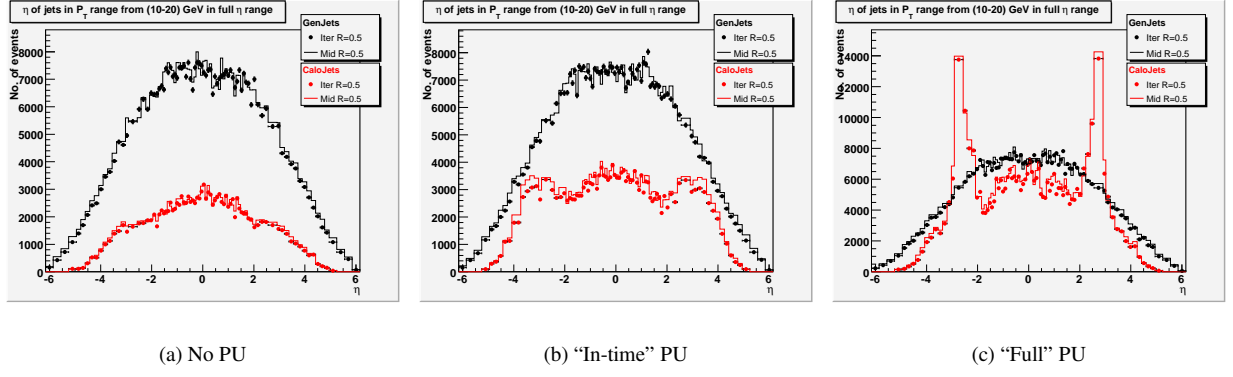


Figure 9: The pseudorapidity(η) distribution of IC5 and MC5 jets with $10 < p_T < 20$ GeV in $50 < \hat{p}_T < 80$ GeV samples for different cases of PU, for Calo and Gen jets.

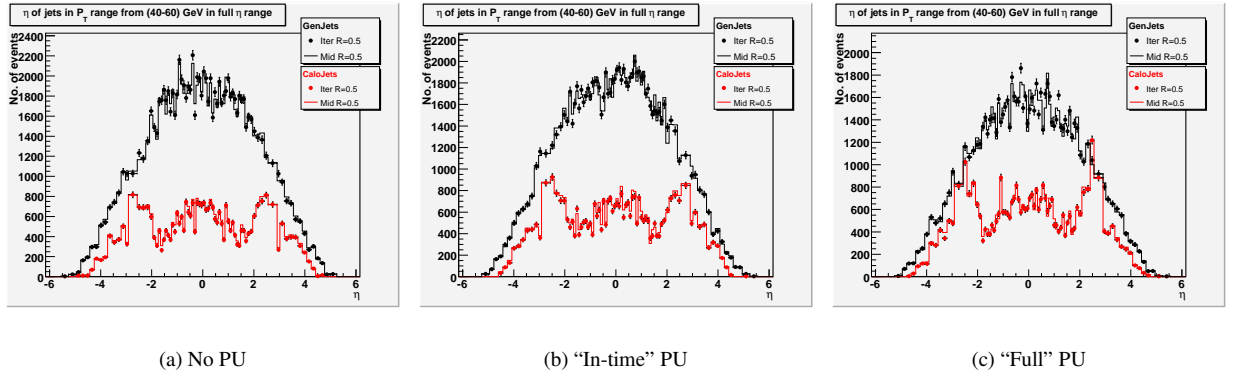


Figure 10: The pseudorapidity distribution of IC5 and MC5 jets with $40 < p_T < 60$ GeV in $50 < \hat{p}_T < 80$ GeV samples for different cases of PU, for Calo and Gen jets.

6.1 Analysis

For the study described in this section, the matching between Calo and Gen jets had to satisfy the restriction that $\Delta R_{min} < 0.25$, or else the candidate match was discarded. In addition, a minimum generated transverse energy cut of 10 GeV was imposed. In the case that a calorimeter jet was successfully matched with a generated jet, the jet response, $R_{jet} = E_T^{REC} / E_T^{GEN}$, was calculated.

Three generated pseudorapidity bins were defined, each meant to highlight a distinct region of the detector. The first, designated as a sufficient representation of the barrel region, includes jets with pseudorapidity between $0 < |\eta_{GEN}| < 1.4$. For the endcap region, the pseudorapidity range of $1.4 < |\eta_{GEN}| < 3.0$ is used, and for the forward region, the pseudorapidity range $3.0 < |\eta_{GEN}| < 5.0$ is used.

6.2 Jet Response

For the results discussed in this section, the histograms of the response were fit with a Gaussian function in the interval $\pm 1.5\sigma$ centered about the mean. The mean parameters of the Gaussian fits were extracted and used as data points to determine the jet response curve as a function of E_T^{GEN} . The results are shown in Fig. 11 for the IC5 algorithm for samples with no PU and in-time PU. The effect of in-time PU is observed only at the low end of the E_T range.

Again, these curves illustrate how the jet response increases with increasing transverse energy. The jet response

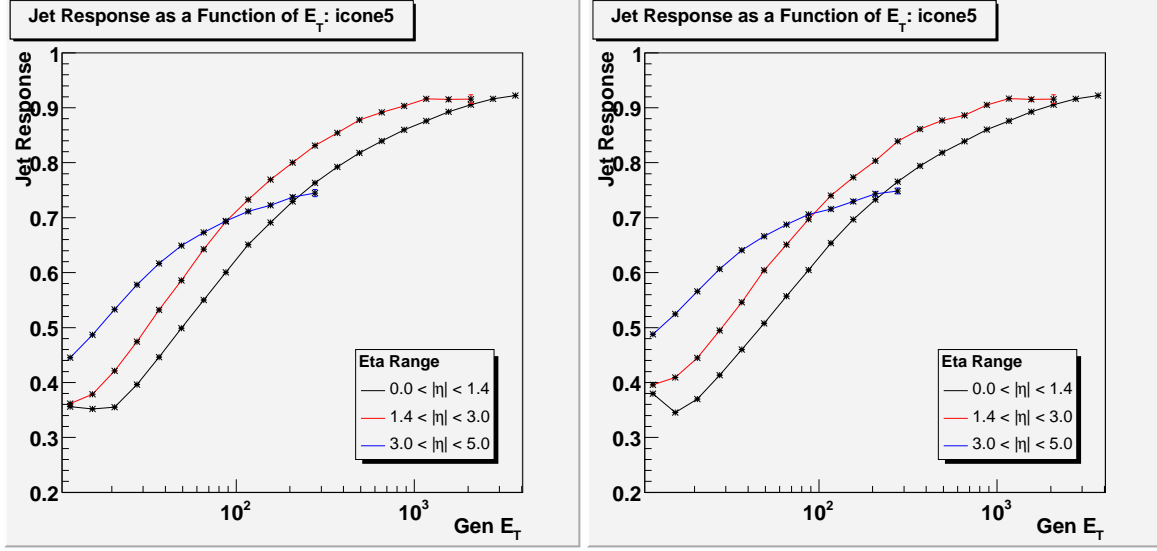


Figure 11: The jet response as a function of generated transverse energy, as reconstructed by the iterative cone algorithm with $R = 0.5$, for three detector regions and samples with no PU (left) and for the case of in-time PU (right).

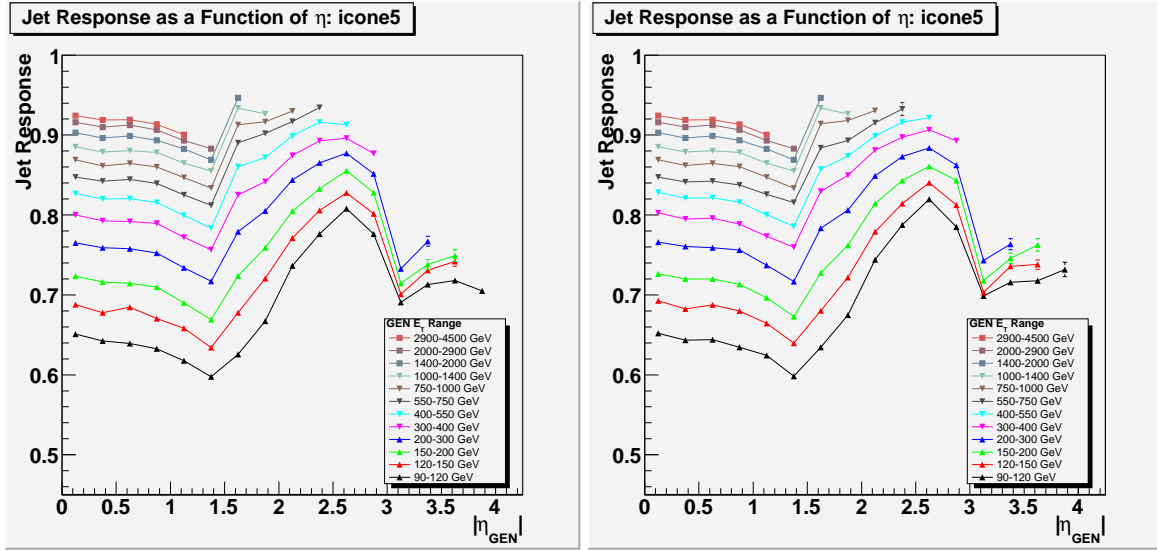


Figure 12: The jet response as a function of generated pseudorapidity, as reconstructed by the iterative cone algorithm with $R = 0.5$, in generated transverse energy bins for samples with no PU (left) and for the case of in-time PU (right).

in the barrel and the endcap levels off between 90% and 95% for jets with generated transverse energy of over 1 TeV. The response in the forward region saturates much sooner, for jets with transverse energy of about 200 GeV at about 75%, as discussed before.

A similar approach was taken when plotting the jet response as a function of generated η . The range of E_T^{GEN} was partitioned into 12 bins, covering a span of $90 \text{ GeV} < E_T^{GEN} < 4500 \text{ GeV}$, and plotted individually. Figure 12 shows the jet response curve as a function of $|\eta_{GEN}|$ for cases with no PU and in-time PU. No significant impact of PU is observed for jet $E_T > 90 \text{ GeV}$. The dip of the jet response curves at $|\eta| = 1.4$ and $|\eta| = 3.0$ corresponds to the transition region of the detector.

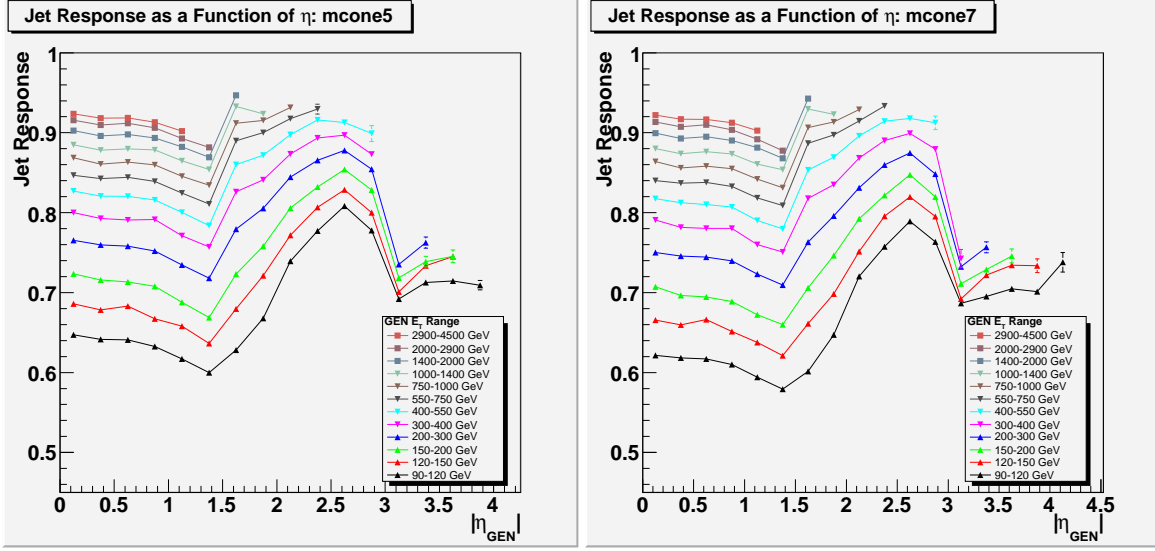


Figure 13: The jet response as a function of generated pseudorapidity, as reconstructed by the MC5 (left) and MC7 (right) algorithms with $R = 0.5$, in generated transverse energy bins for samples with no PU.

The jet response study was repeated using the MC5 and MC7 reconstruction algorithms, as illustrated in Fig. 13.

6.3 Jet Resolutions

The jet resolutions have been calculated for jet energies as well as for jet angular positions (η and ϕ). The results were separated according to detector region and plotted as a function of E_T^{GEN} . The plots were then fitted with a function of the following form:

$$\sqrt{\left(\frac{a}{E_T^{GEN}}\right)^2 + \left(\frac{b}{\sqrt{E_T^{GEN}}}\right)^2 + c^2} \quad (8)$$

where a , b , and c are the fit parameters; they are displayed in the legends for some of the figures. The plots were compared to those obtained with ORCA.

6.3.1 Resolution of the Jet Response

Using the histograms of the jet response, the fractional jet E_T resolution

$$\frac{\sigma(R_{jet})}{\langle R_{jet} \rangle} = \frac{\sigma\left(\frac{E_T^{REC}}{E_T^{GEN}}\right)}{\left\langle \frac{E_T^{REC}}{E_T^{GEN}} \right\rangle} \quad (9)$$

was calculated as a function of E_T^{GEN} , using Gaussian fits (for these fits, the full range was fitted with a Gaussian). For the IC5 algorithm, the resolution curves before and after the application of the jet energy corrections for the three detector regions are displayed in Fig. 14 for the no-PU case. Direct comparisons between uncorrected and corrected resolutions in the three regions of the detector for the IC5 algorithm are shown in Fig. 15. As expected, the jets exhibit an improved resolution when the jet corrections are applied.

The resolutions for the MC5 and MC7 jets after respective energy corrections are shown in Fig. 16.

In Fig. 17 we compare IC5 jet resolutions in the three detector regions derived from samples with in-time and full PU. We conclude that PU effects on the presented resolutions are small over the E_T^{GEN} range shown.

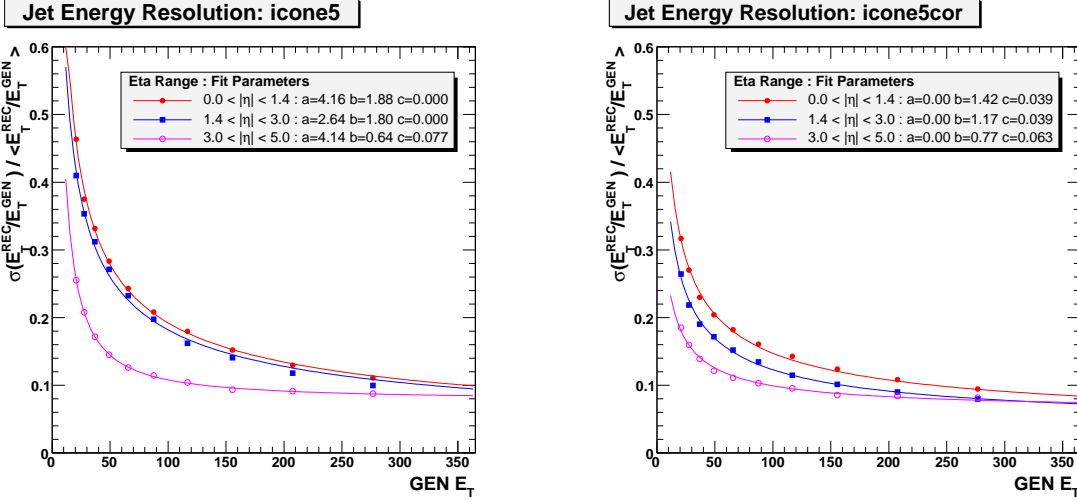


Figure 14: The jet resolution curves (no PU) as functions of E_T^{GEN} , divided by detector region, when reconstructed with the IC5 algorithm. (left) Resolution prior to the application of the jet corrections. (right) Resolution after the jet corrections have been applied.

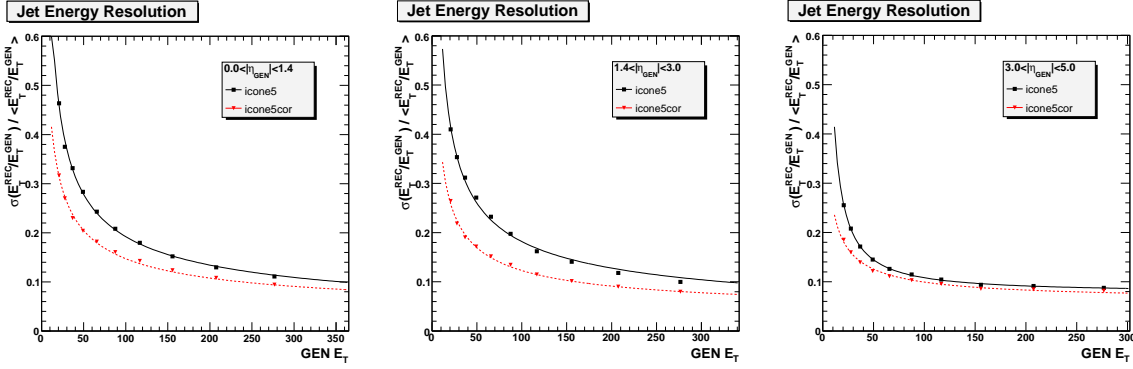


Figure 15: The comparisons of the IC5 jet resolutions before and after jet energy correction (for no PU) as functions of E_T^{GEN} for the three detector regions.

The resolution plots were compared to ORCA as part of the validation study. The comparisons of both corrected and uncorrected jets are shown in Fig. 18. Overall, the resolution curves of CMSSW match relatively well to those of ORCA.

6.3.2 Resolutions of Jet Positions in ϕ and η

The resolutions of jet positions in ϕ and η were calculated from distributions of the following variables for the matched jets:

$$\Delta\phi = \phi_{REC} - \phi_{GEN}, \quad (10)$$

$$\Delta\eta = |\eta_{REC}| - |\eta_{GEN}|, \quad (11)$$

and taking care of the $\pm\pi$ wrap-around in ϕ . We use the absolute values of generated and reconstructed jet η to eliminate a reconstruction bias that pulls the calorimeter jet towards $\eta = 0$ relative to the generated jet [12]. This effect visibly affects the results in HF. Figure 19 shows the resolutions in ϕ and η . Both are displayed as a function of E_T^{GEN} for the IC5 algorithm.

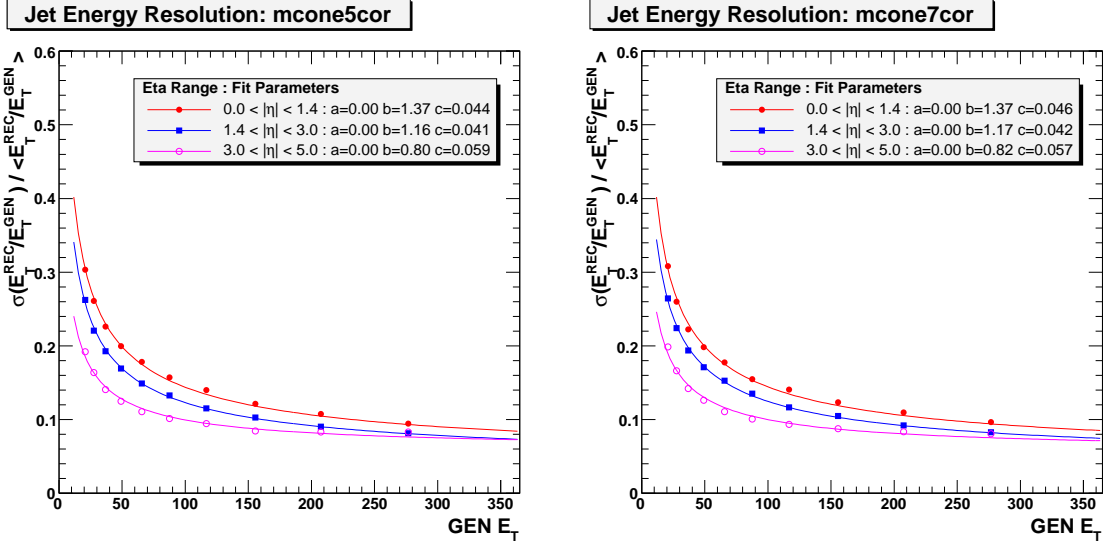


Figure 16: The jet energy resolution as a function of generated transverse energy, when reconstructed with the MC5 (left) and MC7 (right) algorithms in CMSSW_1.2.0, for each of three detector regions, after respective energy corrections and for no-PU conditions.

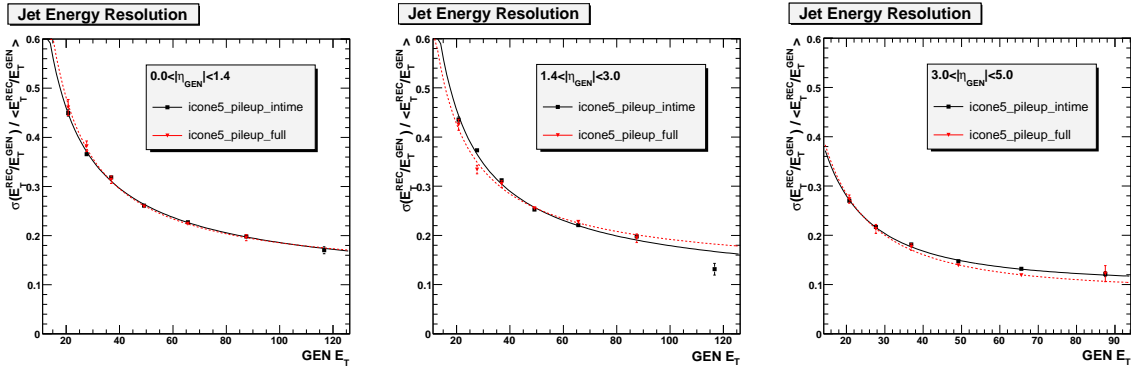


Figure 17: Comparison of IC5 jet resolutions for in-time (black squares) and full (red triangles) pileup conditions in three regions of the detector. No energy corrections have been applied.

The jet position resolutions in HF (for both $\Delta\phi$ and $\Delta\eta$) in CMSSW_1.2.0 are by a factor of 2 worse than previously determined in ORCA. This has been traced back to the Shower Library in HF and improvements have been implemented in CMSSW_1.2.2 and higher versions. For the validation exercise, we used a special small-scale private simulation based on CMSSW_1.2.2 to test the effect of the fix. These results are presented in Fig. 20. A large statistics result from the Spring07 samples using CMSSW_1.2.3 simulation and 1.3.1 reconstruction has become available recently and is presented in Fig. 21.

7 Jet Efficiencies

Jet finding efficiencies were determined for the iterative and midpoint cone jet algorithms for cone sizes of $R = 0.5$ and $R = 0.7$. Jets reconstructed using the k_T algorithm were omitted from this study because of implementation problems in CMSSW_1.2.0. The jet finding efficiency was defined as the ratio of the number of generated jets that matched a calorimeter jet to the total number of generated jets. The same jet definition is used for both calorimeter jets and generated jets. A generated jet was considered to be matched if the nearest calorimeter jet was within a distance $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} = 0.3$. Distributions in ΔR for the four jet definitions are shown in Fig. 22 for two

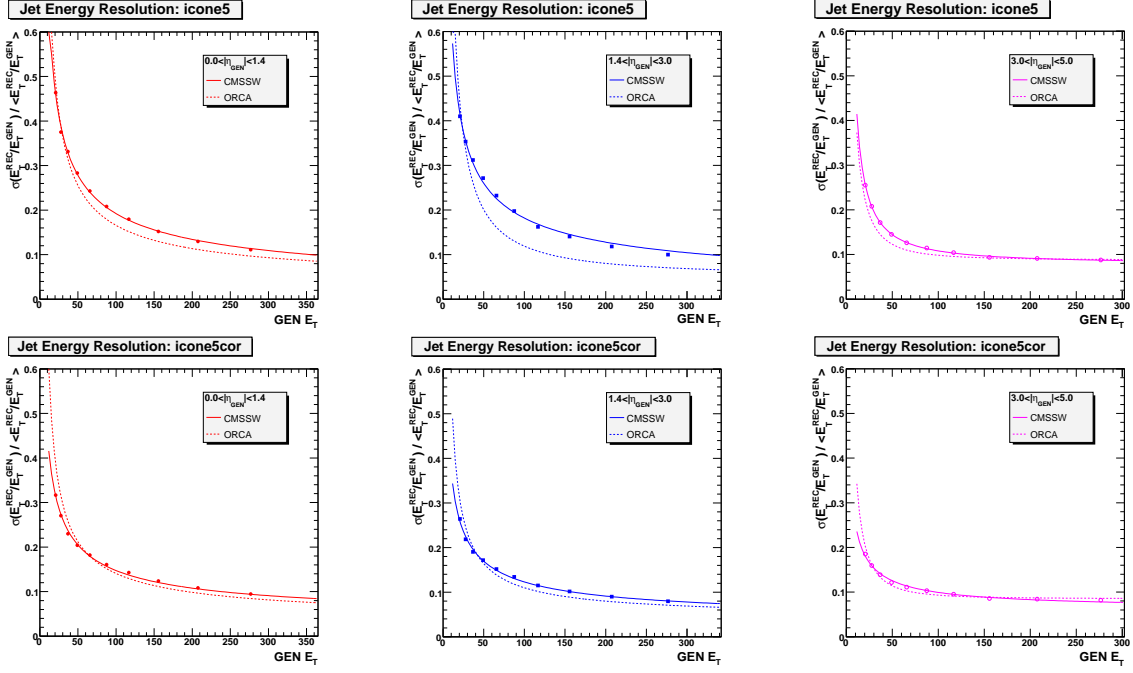


Figure 18: The jet resolution curves of CMSSW (for no PU) compared to ORCA, separated by detector region. Both the uncorrected (top row) and corrected (bottom row) resolutions are shown.

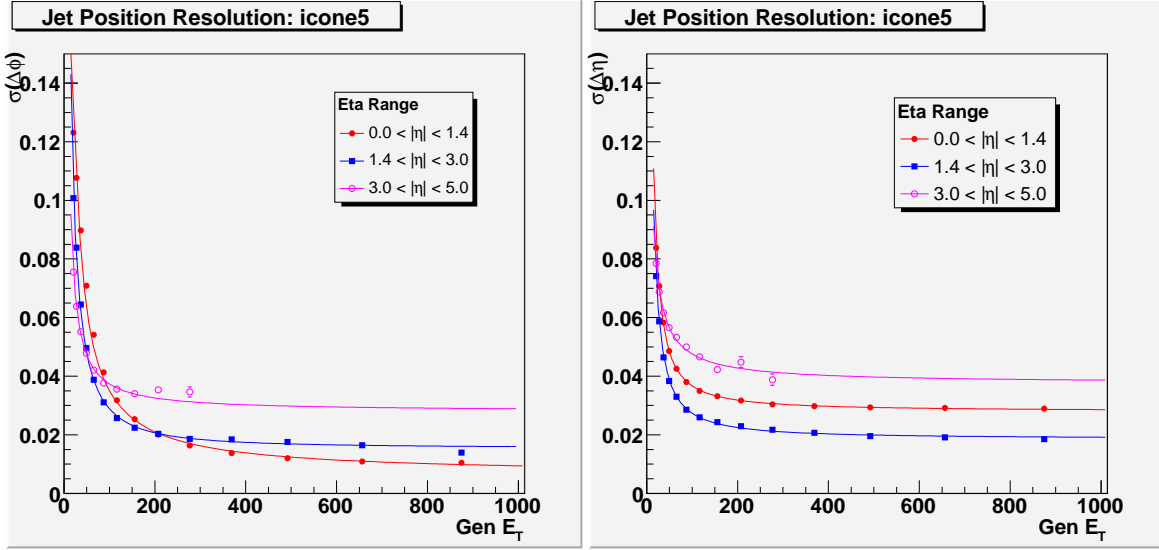


Figure 19: The ϕ (left) and η (right) resolutions as a function of generated transverse energy, when reconstructed with the IC5 algorithm in CMSSW_1.2.0, for each of three detector regions.

\hat{p}_T bins.

Figure 23 shows the jet finding efficiencies as a function of generated jet p_T for $|\eta| \leq 1.5$. There is a relatively large difference in the jet finding efficiencies for the different jet definitions. However, one should note that when comparing the jet finding efficiencies as defined above, the energies assigned to the generated jets can be different for each jet definition. The effects of the different generated jet energies can be taken out by instead using a single jet definition for the generated jets, which provides a common p_T scale for the comparisons, and then finding the number of matched calorimeter jets using each of the different jet definitions. Defined in this manner, the

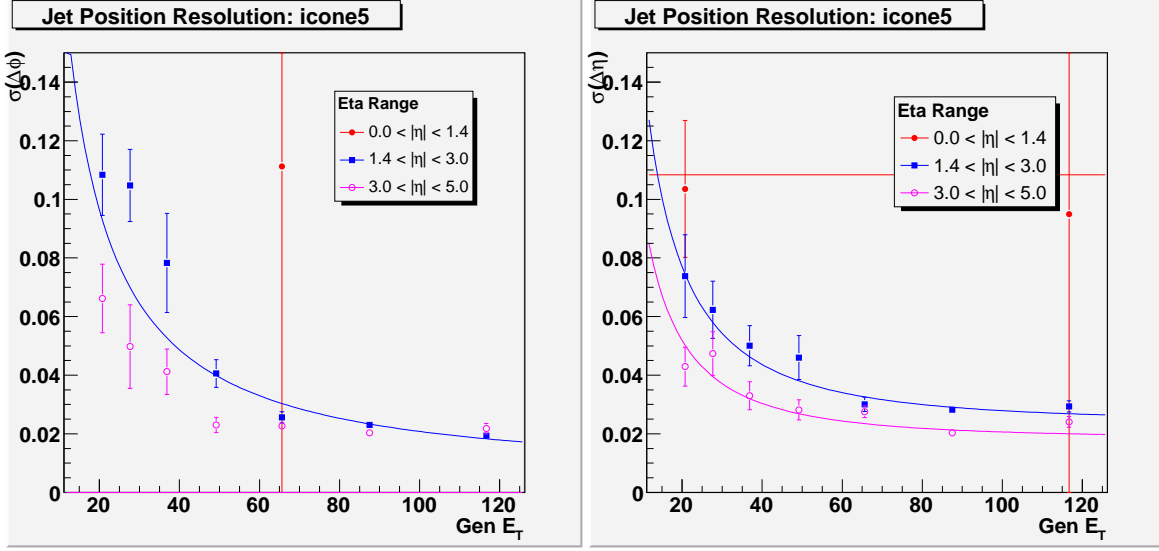


Figure 20: Results of a small scale test of the HF Shower Library fix in CMSSW_1_2_2. The ϕ (left) and η (right) resolutions as a function of generated transverse energy, when simulated with CMSSW_1_2_2 and reconstructed with the IC5 algorithm.

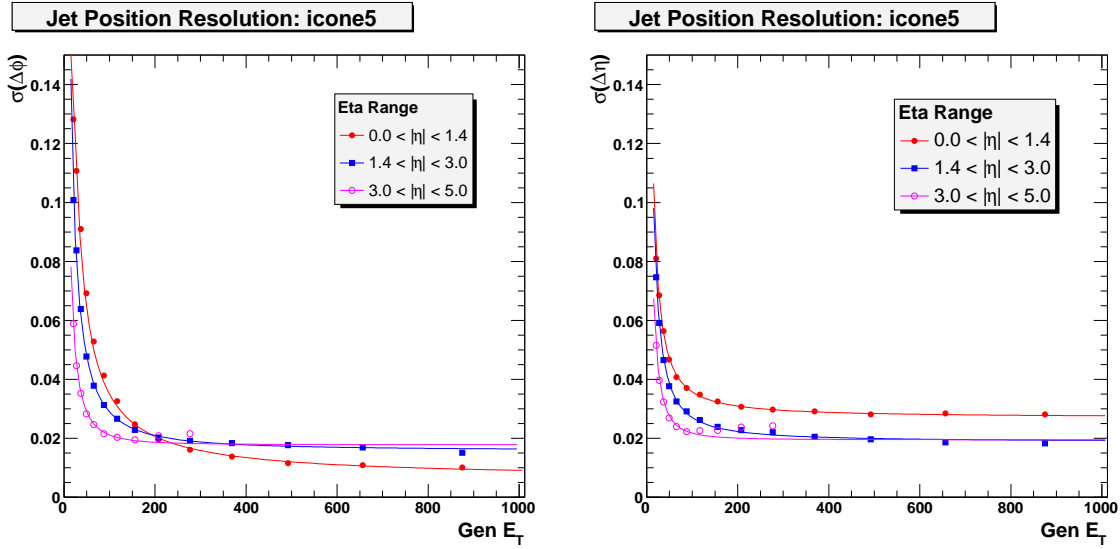


Figure 21: The ϕ (left) and η (right) resolutions as a function of generated transverse energy, when reconstructed with the IC5 algorithm in Spring07 samples, for each of three detector regions.

differences in the jet matching efficiencies are much less pronounced, as shown in Fig. 24. We emphasize that the quantity shown in Fig. 24 does not represent the jet finding efficiency but is only meant to help understand the behavior of the efficiency for different algorithms.

The impact of a minimal p_T requirement for calorimeter jets on the jet finding efficiency is shown in Fig. 25. For the chosen minimal p_T requirements, the effects are small for generated p_T values above 40 GeV. For the minimal p_T requirement of 6 GeV, the jet finding efficiency is $\approx 50\%$ at generated $p_T = 20$ GeV, while for no p_T requirement, the jet finding efficiency is $\approx 50\%$ at generated $p_T = 10$ GeV.

The effects of PU on the jet finding efficiency were also investigated. In Fig. 26, the jet finding efficiency for iterative cone jets with $R = 0.5$ is shown for no PU, in-time PU only, and full PU scenarios. The inclusion of

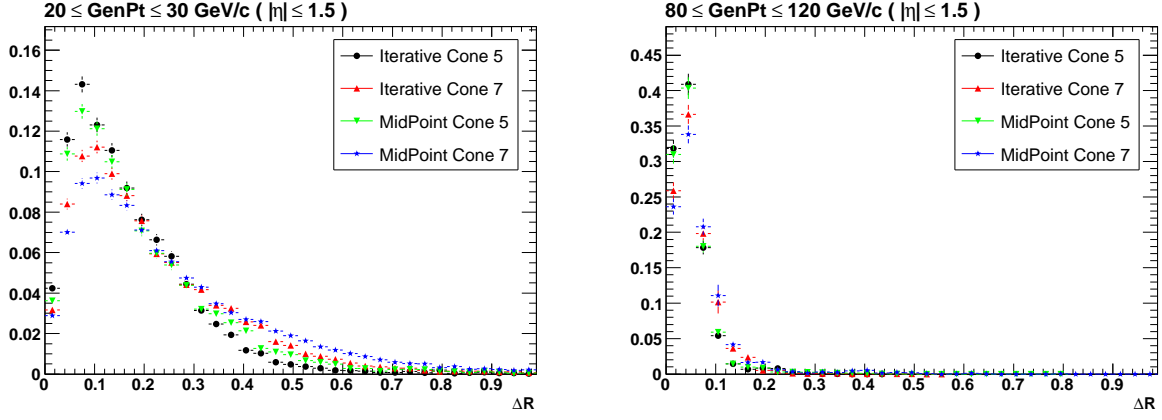


Figure 22: ΔR distributions for the jet p_T bin of 20 – 30 GeV (left) and 80 – 120 GeV (right).

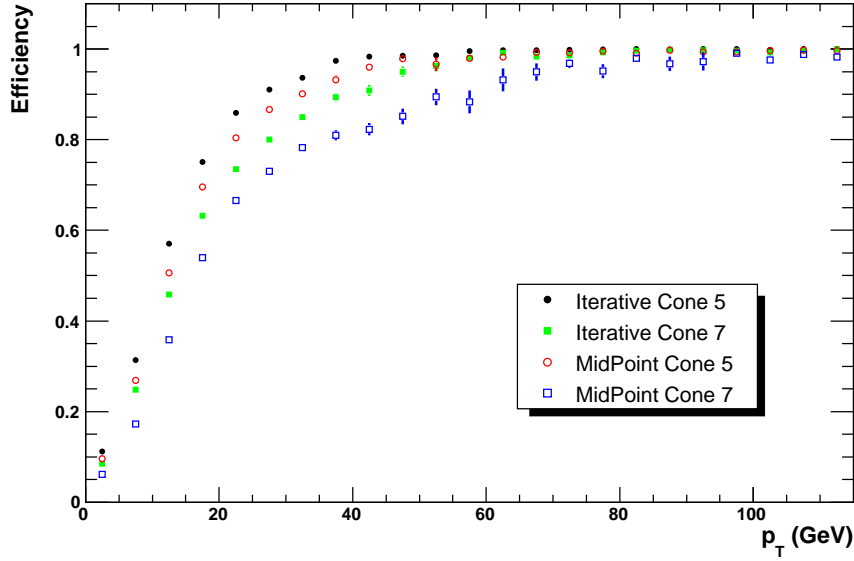


Figure 23: Jet finding efficiency as a function of generated jet p_T for the central calorimeter ($|\eta| \leq 1.5$) and a matching cone of $\Delta R = 0.3$.

extra PU interactions increases the jet finding efficiency at low p_T since these interactions occasionally add enough energy to the calorimeter towers to allow them to pass the jet threshold requirements. (Note that PU particles are not included in generated jets.)

Finally, the jet finding efficiency as function of η is shown in Fig. 27 for two ranges of p_T : $15 < p_T < 25$ GeV and $30 < p_T < 40$ GeV using the samples from Spring07 production. Relatively large differences in the jet finding efficiencies for the different jet definitions are seen for these lower p_T ranges; for $p_T > 100$ GeV all efficiencies are practically flat at 1.0 across the full rapidity range (not shown).

8 Dijet Balance

In $2 \rightarrow 2$ events, transverse momenta of two jets are equal and this property can be used to scale a jet p_T at a given η , to a jet p_T in a reference η region. This procedure can be used to correct for the variation of the jet energy response as a function of η based on data. For the study described below, we define the reference region to be

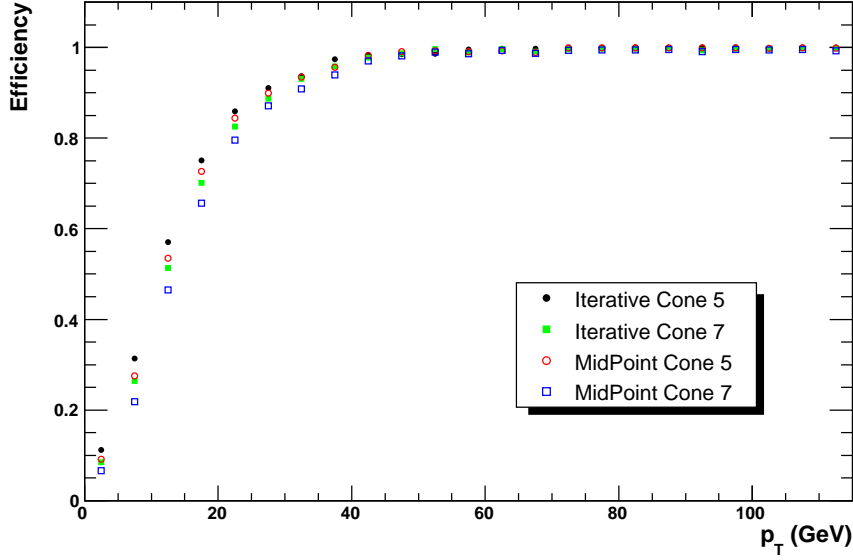


Figure 24: Efficiency of matching calorimeter jets for different jet definitions to generated iterative cone jets, for $R = 0.5$.

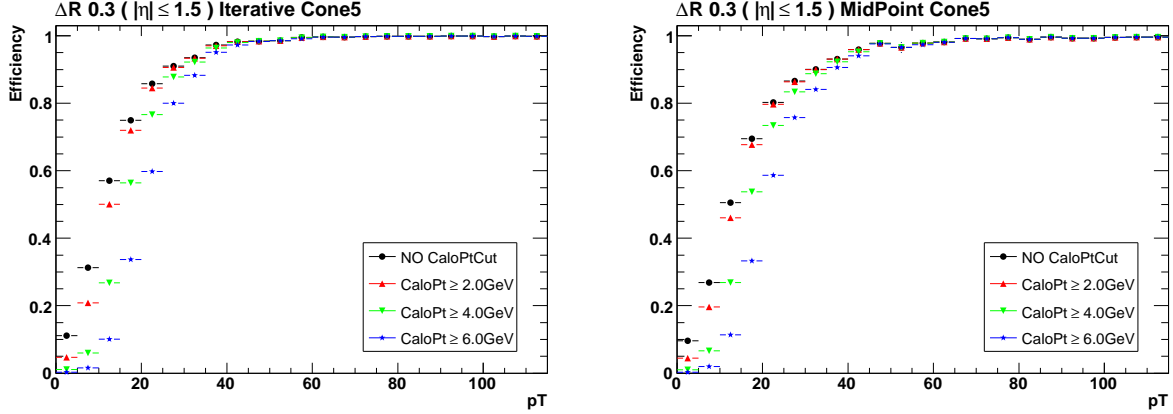


Figure 25: The effects of a minimal calorimeter jet p_T requirement on the jet finding efficiency for IC5 (left) and MC5 (right) jets.

$|\eta| < 1.0$. A similar study was performed using ORCA samples [14].

We define the dijet balance as

$$B(\eta) = \frac{p_T^{Tag} - p_T^{Probe}(\eta)}{(p_T^{Tag} + p_T^{Probe}(\eta))/2} \quad (12)$$

where p_T^{Tag} is the transverse momentum of the jet in the reference region and $p_T^{Probe}(\eta)$ is the transverse momentum of the probe jet at pseudorapidity η . To minimize the effect of jet energy smearing due to finite resolution, the analysis is performed in bins of the average p_T of the two leading jets, $p_T^{ave} = (p_T^{(1)} + p_T^{(2)})/2$.

The relative calorimeter response $R(\eta)$ can then be determined from Eq. 12

$$R(\eta) = \frac{p_T^{Probe}(\eta)}{p_T^{Tag}} = \frac{2 - \langle B(\eta) \rangle}{2 + \langle B(\eta) \rangle} \quad (13)$$

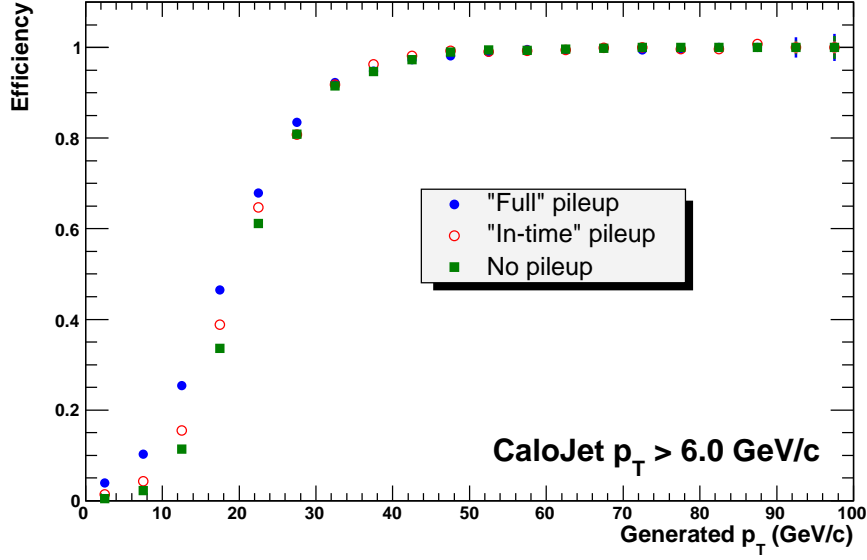


Figure 26: Jet finding efficiency for IC5 jets and various PU scenarios.

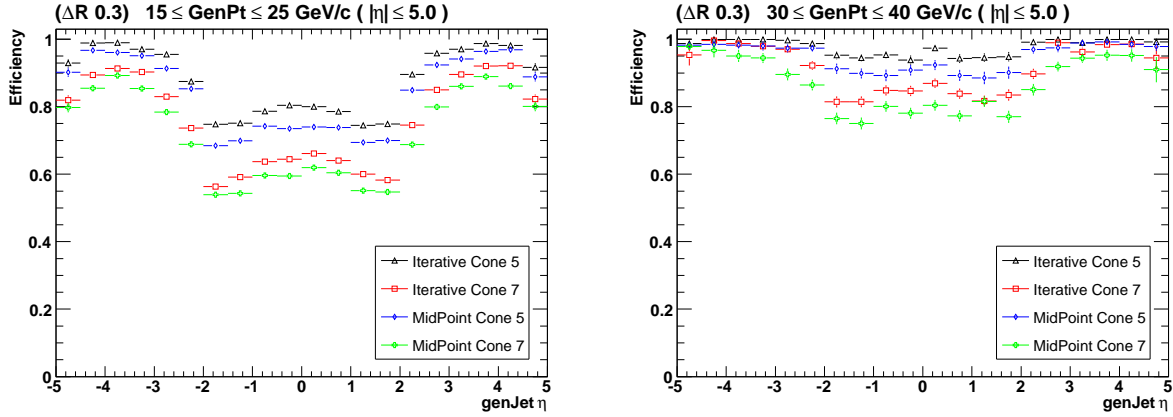


Figure 27: Jet finding efficiency as a function of generated jet η for $15 < p_T < 25$ GeV (left) and $30 < p_T < 40$ GeV (right) using a matching cone of $\Delta R = 0.3$.

where $\langle B(\eta) \rangle$ is the average dijet balance for the probe jets at η . The dijet balance is histogrammed in η bins which have the same size as calorimeter towers. The response, R , is calculated from the mean of the dijet balance histogram. The observable $R(\eta)$ differs from r of Ref. [14] by +1. The measured $R(\eta)$ can be interpolated using a cubic spline and then used to correct the probe jet.

8.1 Analysis

In order to minimize the effect of hard initial and final state radiation (ISR/FSR) which the break $p_T^{(1)} = p_T^{(2)}$ condition, the event is required to satisfy the following requirements:

$$\phi^{(1)} - \phi^{(2)} < 2.7 \quad p_T^{(3)} < \min(10., 0.1 \times p_T^{ave}) \quad (14)$$

where $\phi^{(1)}$ and $\phi^{(2)}$ are the azimuthal angles of the two leading jets and $p_T^{(3)}$ is the transverse momentum of the third leading jet in the event.

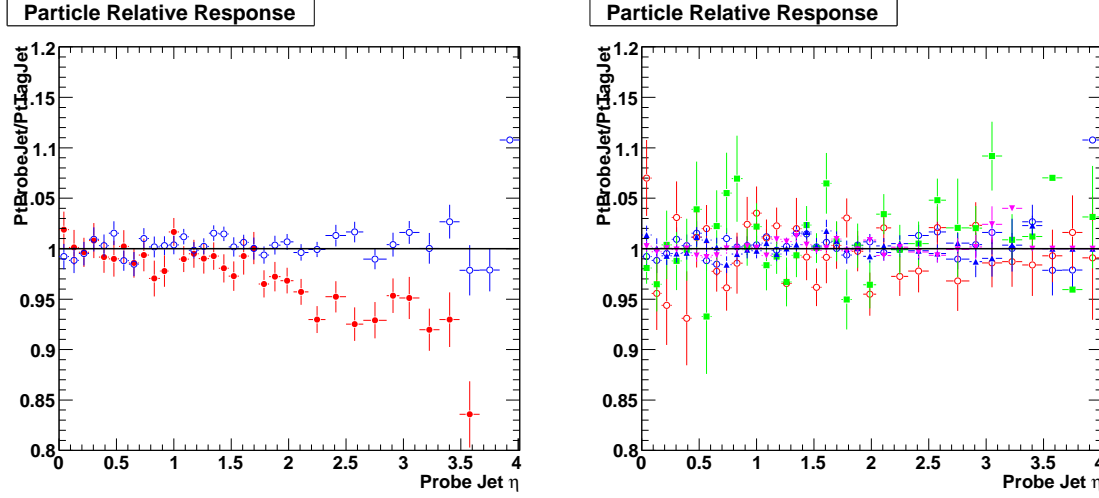


Figure 28: (left) Dijet balancing at the particle level without (filled circles) and with radiation cuts (open circles) for $120 < p_T^{ave} < 230$ GeV; (right) the dijet balancing at the particle level for different p_T^{ave} bins after radiation cuts, Eq. 14.

The impact of the radiation cuts is shown in Fig. 28 for the dijet balance vs. η calculated at the particle level without and with the conditions given by Eq. 14. These results indicate that the concept of dijet balance can be expected to result in a uniform jet energy response when the radiation cuts are applied.

We use the CMSSW_1_2_0 QCD dataset, generated in 21 \hat{p}_T bins. The data are analyzed in various p_T^{ave} bins, weighting events according to production cross section. Figure 29 shows the average p_T^{ave} of two leading jets, their difference in azimuthal angle ($\Delta\phi$), and the transverse momentum of the third jet before and after the cuts. These cuts to suppress radiation were not used in [14] for the relative response study.

The relative response for $120 \leq p_T^{ave} < 230$ GeV bin using raw calorimeter jets is shown in Fig. 30 (left), with and without cuts in Eq. 14. The filled circles show the response without $\Delta\phi$ and third jet cuts and can be compared to Fig. 5 in Ref. [14].

The response for $120 \leq p_T^{ave} < 230$ GeV range, using corrected calorimeter jets is shown in Fig.30 (right); the jets have been corrected using 1_2_0 MCJet Corrections. The filled (open) circles show the response before (after) the extra radiation cuts. As expected, the response is close to unity, except possibly for $|\eta| > 3.5$ region.

The calorimeter relative response may vary with the p_T of the jet. The data were divided in different p_T^{ave} bins (30-60, 60-90, 90-120, 120-250, 250-300, 300-500, 500-800, 800-1000, 1000-2000 GeV). The response for the first six bins is shown in Fig. 31 (left) and for the last four bins in Fig. 31 (right). A variation of $\sim 10\%$ is observed over 30-2000 GeV p_T range, though the statistical uncertainties are large.

9 MET Performance

The Missing Transverse Energy (MET) is determined in CMSSW_1_2_0 from the tranverse vector sum over uncorrected, projective Calorimeter Towers:

$$\mathbf{E}_T^{\text{miss}} = - \sum_n (E_n \sin \theta_n \cos \phi_n \hat{\mathbf{i}} + E_n \sin \theta_n \sin \phi_n \hat{\mathbf{j}}) = -E_x^{\text{miss}} \hat{\mathbf{i}} - E_y^{\text{miss}} \hat{\mathbf{j}} \quad (15)$$

where the index n runs over all calorimeter towers.

In order to conform to the same conditions as the Physics TDR Vol. I [1], or as similar as possible, the default CMSSW “Scheme-B” thresholds were removed, and the Calorimeter Towers were re-reconstructed applying so-called “ORCA PTDR” thresholds, which were simply 0.5 GeV on the energy for all HCAL cells.

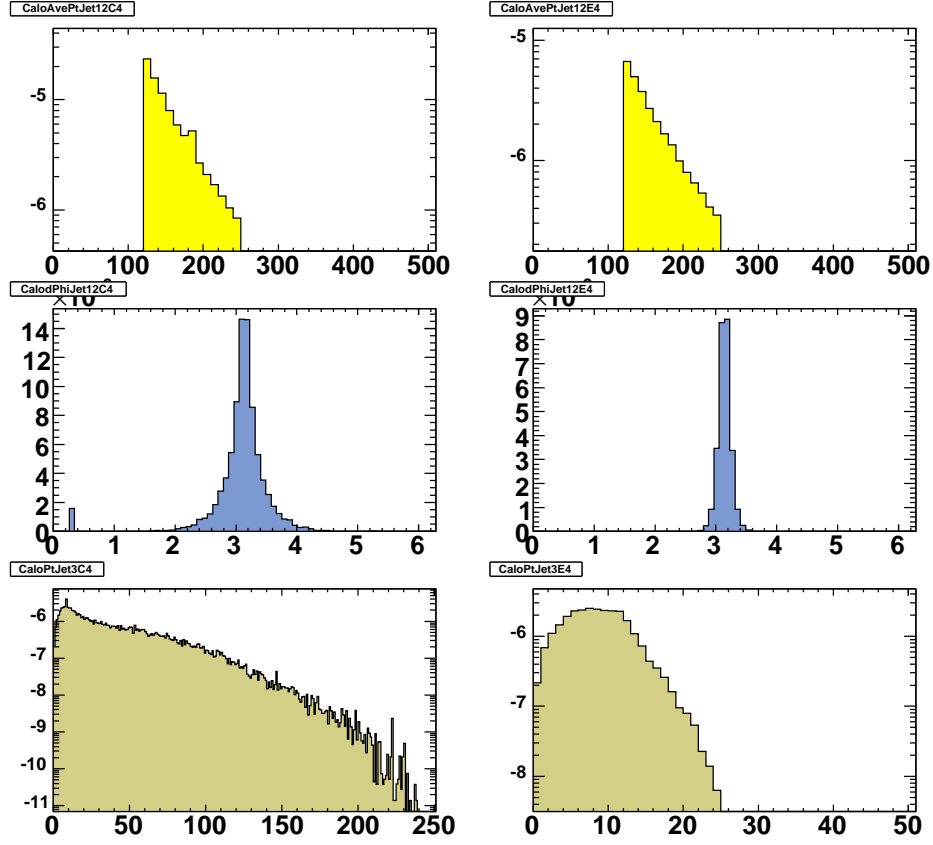


Figure 29: Average p_T , difference in azimuthal angles of the two leading jets and p_T of the third jet before (left) and after (right) applying cuts.

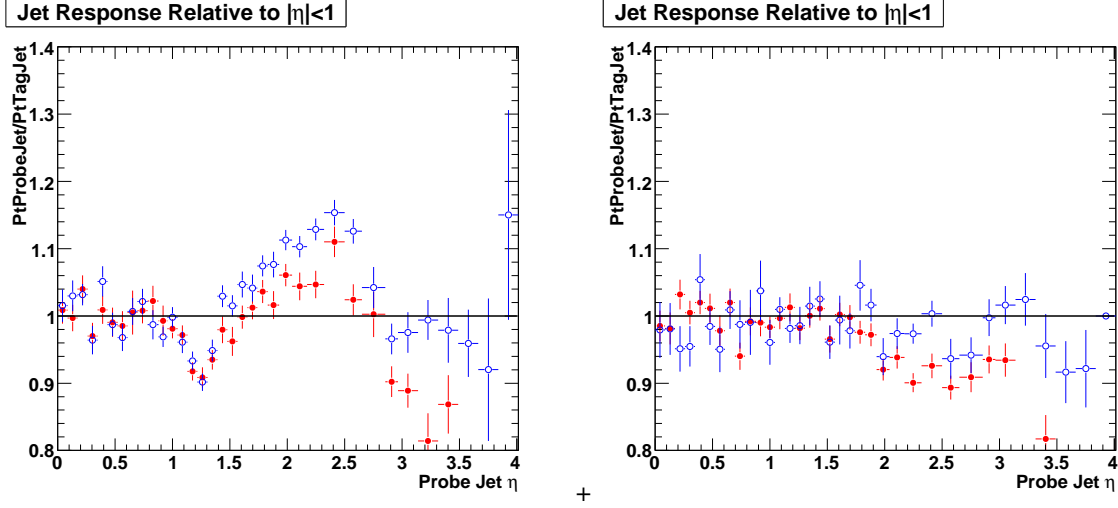


Figure 30: (left) Calorimeter jet response relative to $|\eta| < 1$ region using raw jets before (filled circles) and after (open circles) radiation cuts, Eq. 14. (right) Calorimeter jet response relative to $|\eta| < 1$ region using MCJet-corrected jets before (filled circles) and after (open circles) cuts. Results shown are for $120 < p_T^{ave} < 230$ GeV.

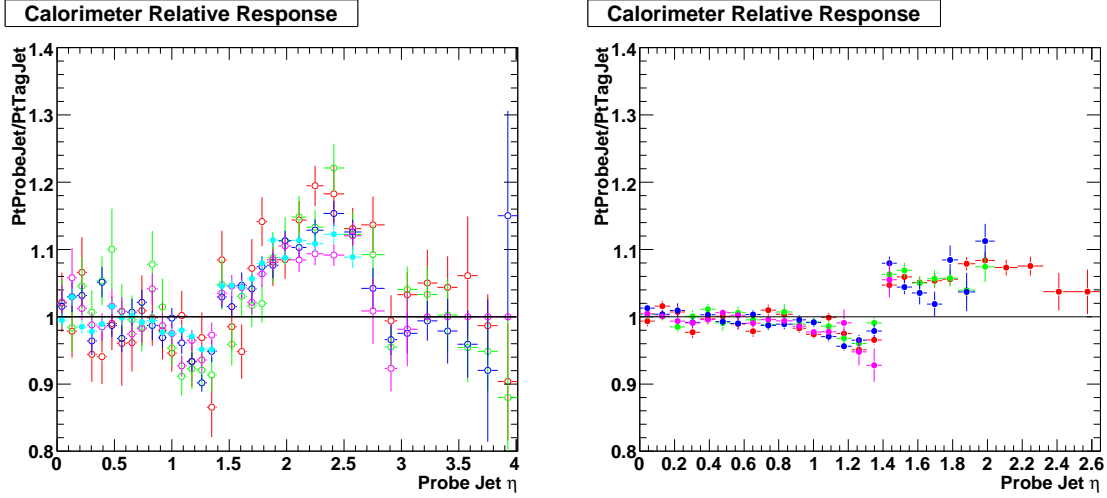


Figure 31: Calorimeter jet response relative to $|\eta| < 1$ region for six p_T^{ave} bins in the 30–500 GeV range (left), and four bins in the 500–2000 GeV range (right); see text for bin values.

The data samples used to study the MET performance correspond to QCD dijets produced in eleven bins of \hat{p}_T between 0 GeV and not more than 800 GeV. In-time PU corresponding to low luminosity running ($2 \times 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$) was also included by default. The effects of out-of-time (integrated over -3 to $+5$ bunch crossings) PU were separately studied.

MET is a simple, global observable which depends on precise symmetric cancelations and hence is a good monitor of the underlying input. The global nature of MET also means that disentangling and understanding the different factors affecting the performance can be very challenging and sometimes impossible. As such, this study attempts to quantify the status of CMSSW_1.2.0 with respect to the ORCA PTDR results and to list, as far as known, the possible differences between the conditions used in the ORCA studies and the conditions used in this study based on CMSSW_1.2.0.

For example, Fig. 32 illustrates the sensitivity of the total visible transverse energy, ΣE_T , to different running and

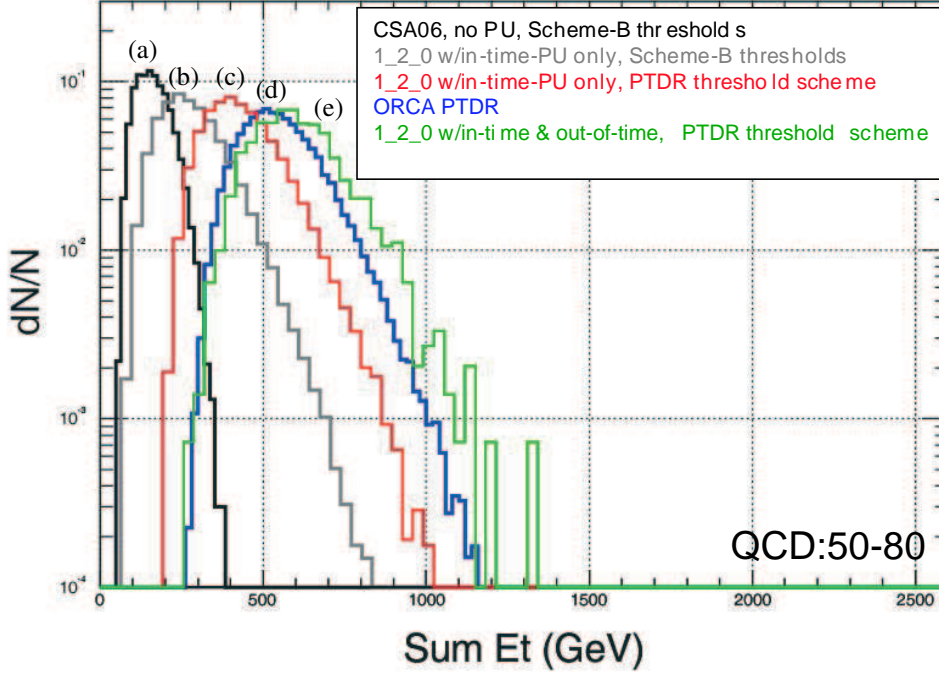


Figure 32: ΣE_T distribution of QCD dijet events with $50 < \hat{p}_T < 80$ GeV. Effects of loosening calorimeter tower thresholds and adding PU can be seen. The result obtained in Ref. [1] using ORCA is seen as the blue histogram (d). The green curve, histogram (e), is obtained from CMSSW_1.2.0 by using ORCA thresholds and includes both in-time and out-of-time PU. See the text for a description of the other histograms (a), (b), and (c).

detector conditions for QCD dijet events with $50 < \hat{p}_T < 80$ GeV. The black histogram (a) represents conditions when there is no PU and when so-called “Scheme-B” thresholds are applied. (“Scheme-B” thresholds are defined in Table 1.) The gray histogram (b) represents the case when in-time PU is included, but “Scheme-B” thresholds are still applied. When one applies the “ORCA PTDR” thresholds together with in-time PU, one obtains the red histogram (c). Finally, when one includes out-of-time PU, together with the “ORCA PTDR” thresholds, one arrives at the green histogram (e). For reference, the results from ORCA [1] are presented as the blue histogram (d). Notice that the final CMSSW_1.2.0 (e) histogram nearly agrees with the ORCA (d) histogram from the Physics TDR, but is slightly shifted to higher energy. This small difference may be due to the fact that in CMSSW the HCAL is read out over more time slices than in ORCA. Indeed, Fig. 32 shows that one must account for a variety of effects in CMSSW (thresholding, PU conditions, read-out configuration, etc) before one can adequately reproduce the ORCA results. This also serves as a warning of the challenges in understanding global quantities, such as the total visible and missing transverse energy, in real data and for the need to exhaustively account for all detector as well as running conditions.

9.1 Missing Transverse Energy in Minimum Bias Events

Figure 33 compares the results presented in Fig. 11.10 of Ref. [1] and the corresponding results of this study for the total visible transverse energy ΣE_T and the x -component of the $\mathbf{E}_T^{\text{miss}}$ (labeled E_x^{miss}) for minimum bias events without PU. As pointed out in the preceding section, the default thresholds applied in ORCA (0.5 GeV for all HCAL cells) are different with respect to the default used in CMSSW (Scheme-B).

While differences between ORCA and CMSSW_1.2.0 remain, replacing the default CMSSW thresholds with the ORCA thresholds recovers much of the behaviour observed in ORCA. In particular, similar shape behaviour for ΣE_T , displayed as plots (a) and (b), is seen between ORCA and CMSSW_1.2.0. An obvious residual difference is observed in the mean value of ΣE_T , which is about 50 GeV lower in CMSSW_1.2.0 compared with ORCA. Also, in the case of E_x^{miss} , displayed as plots (c) and (d), the width of the E_x^{miss} distribution in CMSSW_1.2.0 is

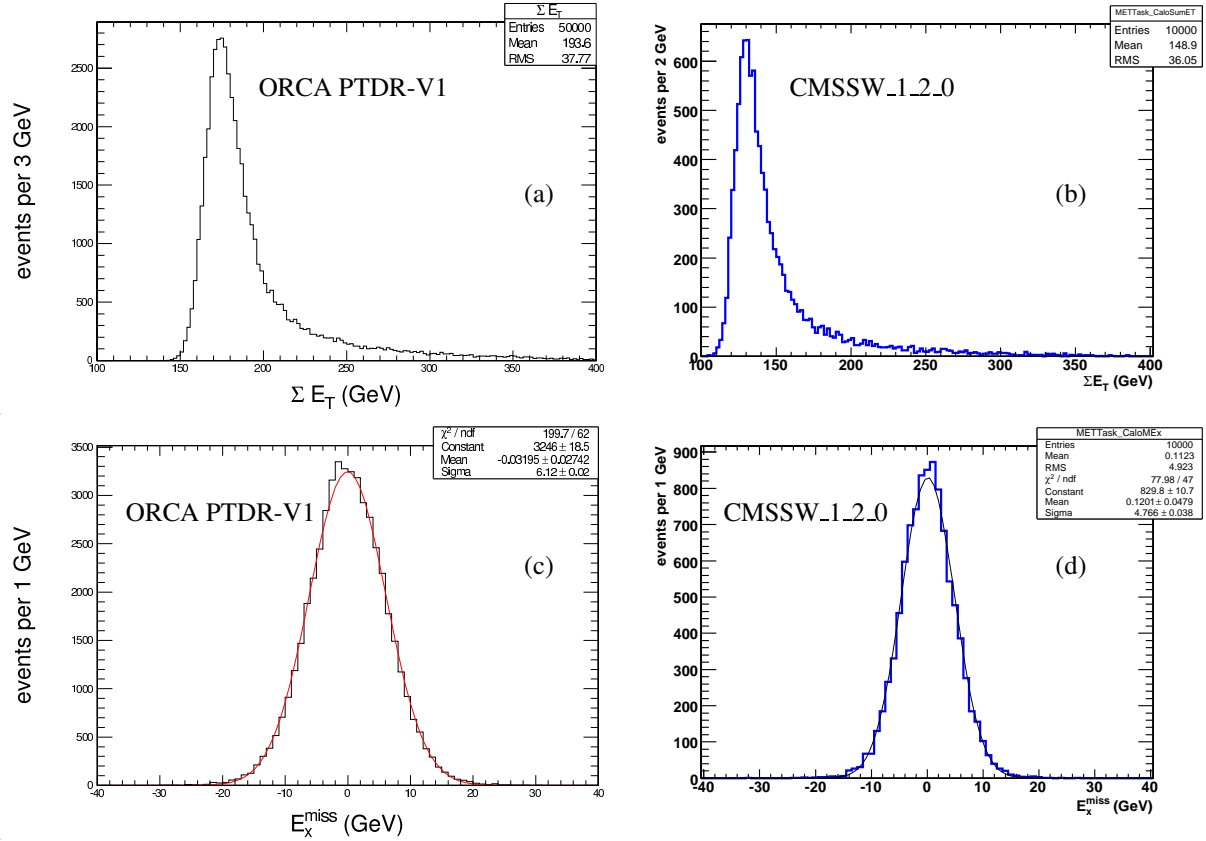


Figure 33: Top row: ΣE_T distributions for minimum bias events without PU in (a) ORCA as presented in Fig. 11.10 of Ref. [1], with a mean of 194 GeV and a RMS of 38 GeV, and (b) from this study using CMSSW_1.2.0, with a mean of 149 GeV and a RMS of 36 GeV. Bottom row: E_x^{miss} distributions for minimum bias events without PU in (c) ORCA as presented in Fig. 11.10 of Ref. [1], with a sigma of 6.1 GeV, and (d) from this study using CMSSW_1.2.0, with a sigma of 4.8 GeV.

observed to be about 20% narrower than in ORCA.

While the origin of these residual differences is not fully understood, we note that there is a known discrepancy in the internal energy scale of the forward hadronic (HF) calorimeter simulation for CMSSW_1.2.0 compared with ORCA (the scale in CMSSW_1.2.0 was improperly set), and that this affects the HF contribution to the MET performance relative to what was in the Physics TDR Vol. I, Ref. [1]. We further note that the HF shower library has been corrected and replaced in CMSSW_1.2.2 and later versions and that the new shower library is both substantially different to that studied in this work and more realistic. Indeed, while the Physics TDR Vol. I results are not expected to be recovered, because of the more realistic HF simulation, comparisons with later versions of CMSSW indicate better agreement with the ORCA results. Nevertheless, a full comparison of the different HF shower libraries is beyond the scope of this note.

9.2 Missing Transverse Energy in Low \hat{p}_T QCD Dijet Events

Figure 34 compares the results presented in Fig. 11.11 of the Physics TDR Vol. I [1] and the corresponding results of this study for ΣE_T and E_x^{miss} for QCD dijet events with $\hat{p}_T < 15$ GeV. Low \hat{p}_T QCD dijet events are similar to minimum bias events; the main difference in this study is that PU corresponding to low-luminosity running are included in the QCD dijet events, whereas in the case of minimum bias events studied in the previous section, no PU effects are included.

As in the previous section, similar shape behaviour for ΣE_T , displayed as plots (a) and (b), is seen between ORCA

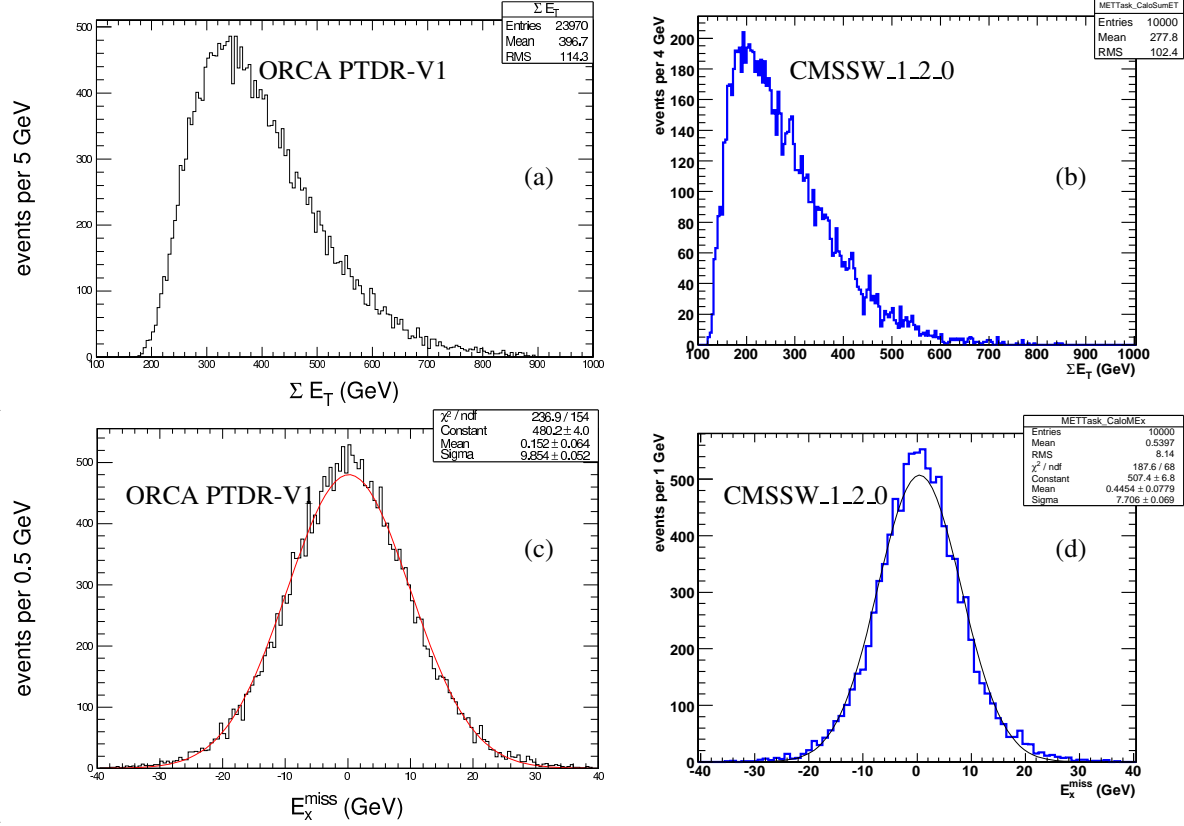


Figure 34: Top row: ΣE_T distributions for QCD dijet events with $\hat{p}_T < 15$ GeV in (a) ORCA as presented in Fig. 11.11 of Ref. [1], with a mean of 397 GeV and a RMS of 114 GeV, and (b) from this study using CMSSW_1.2.0, with a mean of 278 GeV and a RMS of 102 GeV. Bottom row: E_x^{miss} distributions for QCD dijet events with $\hat{p}_T < 15$ GeV in (c) ORCA as presented in Fig. 11.11 of Ref. [1], with a sigma of 9.9 GeV, and (d) from this study using CMSSW_1.2.0, with a sigma of 7.7 GeV.

and CMSSW_1.2.0. However, a shift of more than 100 GeV (with respect to ORCA) to lower average ΣE_T is observed in CMSSW_1.2.0. While PU effects are included, the ORCA studies included both in-time and out-of-time PU, whereas the CMSSW_1.2.0 studies only include in-time PU. Hence, at least part of the shift is understood to be due to the fact that the CMSSW_1.2.0 simulation does not include out-of-time PU. Comparing plots (c) and (d) of Fig. 34, the E_x^{miss} resolution in CMSSW_1.2.0 is observed to be 20% narrower than in ORCA, consistent with the minimum bias results presented in the previous section.

In this study, the performance of MET resolution is parameterised in the total visible transverse energy of the event (ΣE_T) according to the following form:

$$\frac{\sigma(\mathbf{E}_T^{\text{miss}})}{\Sigma E_T} = \frac{a}{\Sigma E_T} \oplus \frac{b}{\sqrt{\Sigma E_T}} \oplus c \quad (16)$$

where the a (“noise”) term represents effects due to electronic noise, PU, underlying event, etc; the b (“stochastic”) term represents the statistical sampling nature of the Calorimeter Towers; and the c (“constant”) term represents residual systematic effects due to non-linearities, cracks, dead material, etc. It is important to emphasize that the above parameterisation factorizes the MET uncertainty into independent effects a , b , c . In particular, the stochastic and constant terms do not depend on the effects due to noise, PU, and underlying event (to first order). Hence, when comparing the MET performance between samples having different noise thresholds and PU conditions (as is done in this study), the offset in ΣE_T must be explicitly taken into account.

Figure 35 illustrates the resolution of the E_x^{miss} (defined to be the root-mean-square of E_x^{miss}) as a function of

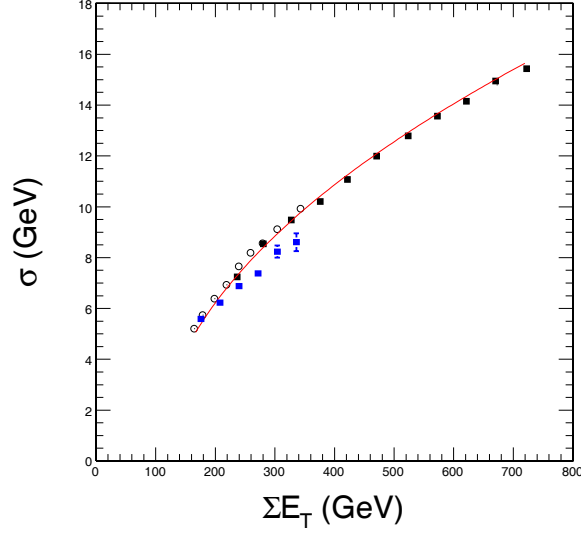


Figure 35: $\sigma(E_x^{\text{miss}})$ vs ΣE_T distribution for QCD dijet events for $\hat{p}_T < 15$ GeV. The square points (blue with error bars) were obtained using CMSSW_1.2.0 and are superimposed on the ORCA results as presented in Fig. 11.12 of Ref. [1] (displayed with white circles, black squares and a fitted red line)

the average ΣE_T for QCD dijet events with $\hat{p}_T < 15$ GeV. The square points (with error bars) were obtained using CMSSW_1.2.0 and are superimposed on the ORCA results (displayed with white circles, black squares and a fitted red line). Notice that while the average ΣE_T and $\sigma(E_x^{\text{miss}})$ distributions are different when comparing CMSSW_1.2.0 with ORCA, the relationship between the two variables in CMSSW_1.2.0 is somewhat similar with respect to ORCA. At such low \hat{p}_T , noise and PU dominate the MET resolution, and hence the effect of out-of-time PU ought to play a role in the low MET resolution: A fit to the CMSSW_1.2.0 data yields a constant term of 1.8% and a stochastic term of 47% $\text{GeV}^{1/2}$, which should be compared with the corresponding result from ORCA reported in the PTDR of 65% $\text{GeV}^{1/2}$ for the stochastic term (no constant term is reported in the PTDR).

9.3 Missing Transverse Energy in All QCD Dijet Events

Figure 36 compares the behaviour of ΣE_T distributions for different QCD dijet samples within $20 < \hat{p}_T < 800$ GeV and between ORCA and CMSSW_1.2.0. Again, the CMSSW_1.2.0 simulation corresponds to in-time PU only, whereas the ORCA simulation includes both in-time and out-of-time PU. Consistent with Fig. 34, the shape of the ΣE_T distributions for ORCA and CMSSW_1.2.0 distributions appear to be similar, but with peak ΣE_T values in CMSSW_1.2.0 shifted 100 GeV to 200 GeV lower compared to ORCA. As demonstrated in Fig. 32, nearly all of this difference can be accounted for by the lack of out-of-time PU in CMSSW_1.2.0.

Figure 37 illustrates the behaviour of $\sigma(E_x^{\text{miss}})$ with ΣE_T for QCD dijet samples within $20 < \hat{p}_T < 800$ GeV. The blue points correspond to CMSSW_1.2.0 and the black points (with corresponding red fitted curve) correspond to ORCA. Both plots (left and right) are the same, except that different ranges of ΣE_T are used in the fit for the CMSSW_1.2.0 points. Also, since we know that: (1) there are differences between ORCA and CMSSW_1.2.0 in the offset of ΣE_T , and (2) we are most interested in comparing the performance of the stochastic and constant terms of the MET resolution, we explicitly include an offset parameter for ΣE_T in the fit. A fit to the full range of available data points (left plot of Fig. 37) yields the function for the MET resolution: $\sigma^2 = (9.487 \text{ GeV})^2 + (89\% \text{ GeV}^{1/2})^2 (\Sigma E_T - 369.4 \text{ GeV}) + (2.3\% (\Sigma E_T - 369.4 \text{ GeV}))^2$. Compared with the Physics TDR Vol. I, where ORCA gave a stochastic term of 97% and a constant term of 1.2%, the fit gives a somewhat smaller stochastic term, but a significantly larger (by nearly a factor of two) constant term. Finally, the goodness of fit corresponds to a χ^2 of 45.3 per 8 degrees of freedom, with the largest residual from the highest ΣE_T points.

The fact that the highest data points for ΣE_T in CMSSW_1.2.0 appear to *possibly* more closely follow the ORCA

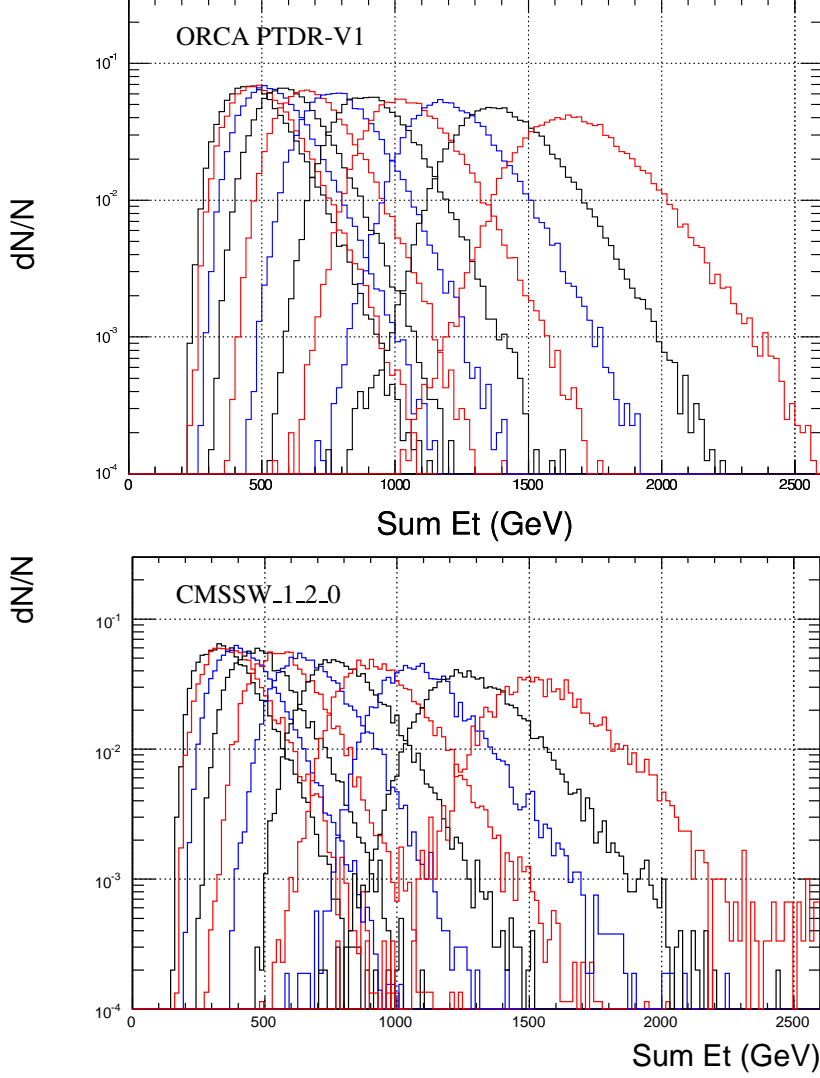


Figure 36: ΣE_T distributions for QCD dijet events for \hat{p}_T ranges 20-30, 30-50, 50-80, 80-120, 120-170, 170-230, 230-300, 300-380, 380-470, 470-600, and 600-800 GeV going from left to right. Top plot: ORCA results as presented in Fig. 11.13 of Ref. [1]. Bottom plot: results from this study using CMSSW_1_2_0.

results, may be a hint that, in CMSSW_1_2_0, the low ΣE_T regime behaves differently from the high ΣE_T regime. To test this hypothesis, two additional fits are performed: one to the low ΣE_T regime (not shown) and another to only the high ΣE_T regime (right plot of Fig. 37). The fit to just the low ΣE_T regime, yields a better fit (χ^2 of 13.8 per 5 degrees of freedom) with very similar stochastic and constant terms as the full fit. The fit to just the high ΣE_T (plot on the right of Fig. 37), is basically inconclusive (χ^2 of 2.9 per 1 degree of freedom), but does seem to hint that the results (a stochastic term of 112% and no constant term) may be more similar to that obtained with ORCA. Table 5 summarises the results of the different fit ranges.

Figure 38 illustrates the E_T^{miss} distributions for QCD dijet samples within $20 < \hat{p}_T < 800$ GeV and having in-time PU only. The peak values for these distributions are generally in good agreement with those shown by ORCA. However, there is a residual difference in that each distribution is somewhat more narrow, resulting in smaller $\text{RMS}(E_T^{\text{miss}})$ values. Such behaviour is consistent with the fact that the CMSSW_1_2_0 samples lack of out-of-time PU, resulting in less transverse energy activity.

Figure 39 shows the dependence of $\langle E_T^{\text{miss}} \rangle$ as a function of ΣE_T for all QCD dijet samples with $\hat{p}_T < 800$

Fit Range	Stochastic	Constant	χ^2/NDF
full	$89\% \pm 1\%$	$2.3\% \pm 0.1\%$	45.3/8
low ΣE_T	$85\% \pm 1\%$	$2.8\% \pm 0.1\%$	13.8/5
high ΣE_T	$118\% \pm 3\%$	$0\% \pm 2\%$	2.9/1
PTDR [1]	97%	1.2%	N/A

Table 5: Summary of fit results for $\sigma(E_x^{\text{miss}})$ vs ΣE_T in QCD dijet samples with $\hat{p}_T < 800$ GeV. For comparison, results of Ref. [1] are also given.

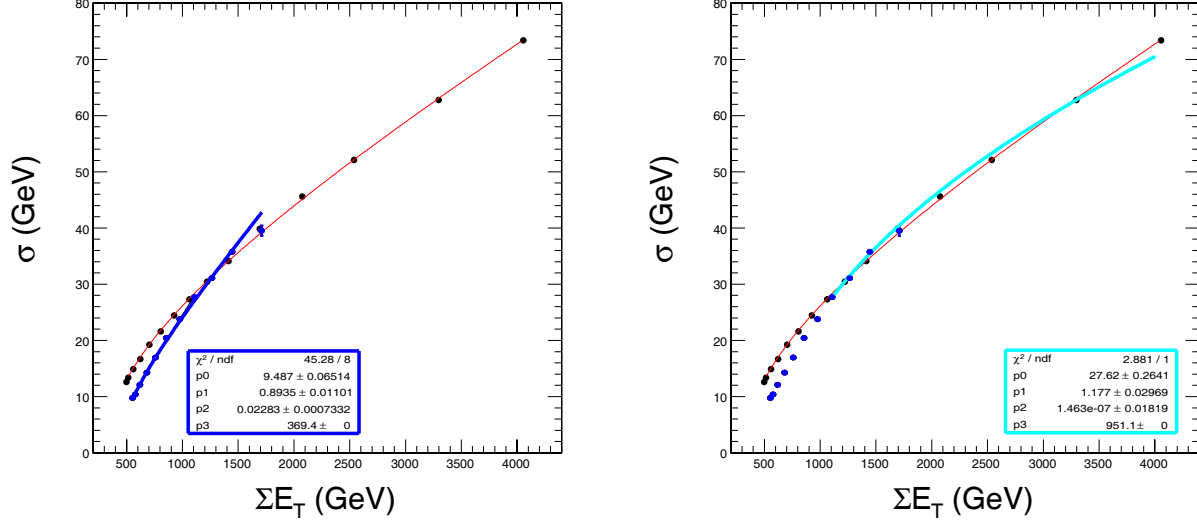


Figure 37: $\sigma(E_x^{\text{miss}})$ vs ΣE_T for QCD dijet samples with $20 < \hat{p}_T < 800$ GeV. The blue points were obtained using CMSSW_1_2_0 and are superimposed on the ORCA results as presented in Fig. 11.14 of Ref. [1] (displayed with black dots and a fitted red line). Left plot: fit to the full range of available data points. Right plot: fit to only the high ΣE_T regime.

GeV. The blue points correspond to CMSSW_1_2_0 and the black points (with corresponding red fitted curve) correspond to ORCA. As in Fig. 37, both plots (left and right) are the same, except that different ranges of ΣE_T are used in the fit for the CMSSW_1_2_0 points. Because there is no out-of-time PU included in the QCD samples, an overall offset in ΣE_T is removed by the fit. A fit to the full range of available data points (left plot of Fig. 39) yields the function $\langle E_T^{\text{miss}} \rangle^2 = (9.745 \text{ GeV})^2 + (83\% \text{ GeV}^{1/2})^2 (\Sigma E_T - 277.8 \text{ GeV}) + (4\% (\Sigma E_T - 277.8 \text{ GeV}))^2$. For comparison, the Physics TDR Vol. I [1] with ORCA obtained a stochastic term of 123% and a constant term of 1.9%.

It is clear from the χ^2 of 232.9 per 10 degrees of freedom, that the functional form does not fit the data points well. Hence, as in Fig. 37, two additional fits are performed: one to the low ΣE_T regime (not shown) and another to only the high ΣE_T regime (right plot of Fig. 39). The fit to the low ΣE_T regime yields a much better fit (χ^2 of 16.1 per 5 degrees of freedom) and results in a constant term of 4% (confirming the full fit results), whereas the stochastic term reduces to 71%. In the high ΣE_T regime, the fit is also somewhat better (χ^2 of 8.6 per 2 degrees of freedom) and the constant term reduces to $(2 \pm 0.3)\%$, consistent with the ORCA results reported in the Physics TDR Vol I, while the stochastic term increases to $(146 \pm 12)\%$, which is somewhat (but not dramatically) higher than that obtained with ORCA. Additional data points were not available to confirm this behavior for higher ΣE_T values. Table 6 summarises the results of the different fit ranges.

Figure 40 shows the $\sigma(\text{MET-perp})$ vs ΣE_T distribution, where MET-perp is the component of MET perpendicular to the jet axis. The jet axis is found by defining the unit vector \hat{n} in the direction of $\vec{p}_{T,jet1} - \vec{p}_{T,jet2}$. The resolution of MET-perp is primarily determined by underlying event, PU, and detector noise rather than by the calorimeter

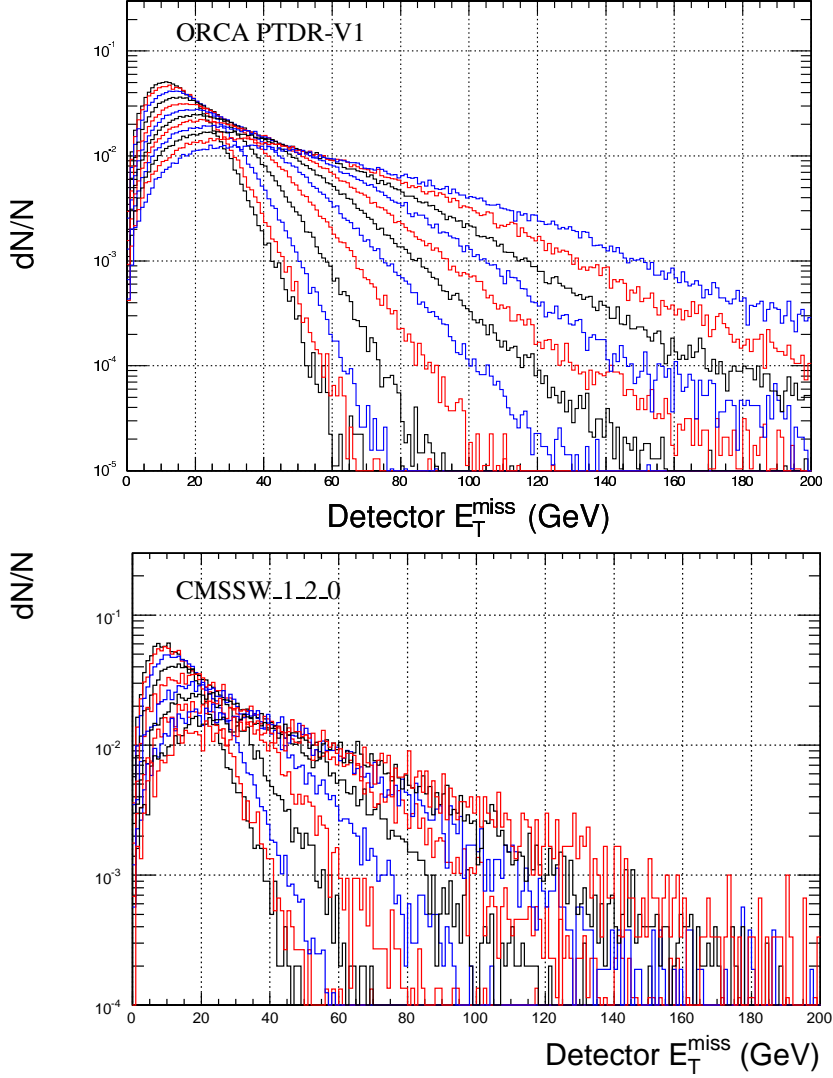


Figure 38: MET distributions for QCD dijet events for \hat{p}_T ranges 20-30, 30-50, 50-80, 80-120, 120-170, 170-230, 230-300, 300-380, 380-470, 470-600, and 600-800 GeV going from left to right. Top plot: ORCA results as presented in Fig. 11.15 of Ref. [1]. Bottom plot: results from this study using CMSSW_1_2_0.

response to the jets. Here we see that CMSSW_1_2_0 is in rough agreement with ORCA.

10 Conclusions

We summarized results of jet and Missing ET performance studies in CMSSW_1_2_0 for the SDPV effort.

Significant changes in the simulation and calibration of calorimeter properties are planned for the CSA07 production. The results presented here will help to evaluate the consequent impact on various aspects of jet and MET performance.

References

- [1] CMS PTDR Vol. I: Detector Performance and Software, CERN/LHCC 2006-001
- [2] SDPV agenda links can be found at: <http://indico.cern.ch/categoryDisplay.py?categId=4155>

Fit Range	Stochastic	Constant	χ^2/NDF
full	$83\% \pm 1\%$	$4.05\% \pm 0.03\%$	232.9/10
low ΣE_T	$71\% \pm 1\%$	$4.9\% \pm 0.1\%$	16.1/5
high ΣE_T	$146\% \pm 13\%$	$2.1\% \pm 0.3\%$	8.6/2
PTDR [1]	123%	1.9%	N/A

Table 6: Summary of fit results for $\langle E_T^{\text{miss}} \rangle$ vs ΣE_T in QCD dijet samples with $\hat{p}_T < 800$ GeV. For comparison, the results of Ref. [1] are also given.

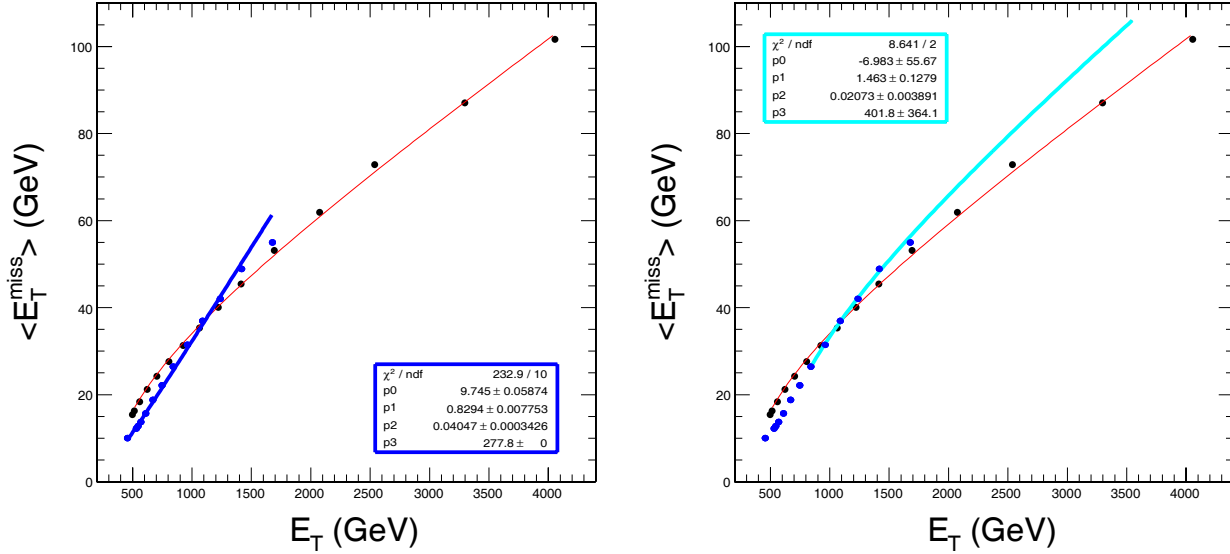


Figure 39: $\langle E_T^{\text{miss}} \rangle$ vs ΣE_T for QCD dijet samples with $\hat{p}_T < 800$ GeV. The blue points were obtained using CMSSW_1.2.0 and are superimposed on the ORCA results as presented in Fig. 11.16 of Ref. [1] (displayed with black dots and a fitted red line). Left plot: fit to the full range of available data points. Right plot: fit to only the high ΣE_T regime.

- [3] SDPV final report is linked from: <http://indico.cern.ch/conferenceDisplay.py?confId=12699>
- [4] R. Demina et al., CMS NOTE-2006/020, “*Calorimeter Cell Energy Thresholds for Jet Reconstruction in CMS*”.
- [5] CMS Note 2006/036, A. Heister et al., “*Measurement of Jets with the CMS Detector at LHC*”.
- [6] Comp. Phys. Comm. vol 153/1 85-96 (2003), J. M. Butterworth, J. P. Couchman, B. E. Cox, B. M. Waugh, “*A C++ implementation of the K_{\perp} clustering algorithm*”.
- [7] M. Cacciari and G. P. Salam, Phys. Lett. B **641**, 57 (2006)
- [8] D. Acosta et al., CMS NOTE-2006/067, “*The Underlying Event at the LHC*”.
- [9] http://www.uscms.org/LPC/lpc_jetmet/edmData/edmData.html
- [10] <https://twiki.cern.ch/twiki/bin/view/CMS/CmsSamples>
- [11] CMS Note 2007/006, The CMS Collaboration, “*CMS Computing, Software and Analysis Challenge in 2006 (CSA06) Summary*”.
- [12] O. Kodolova, ORCA study, private communication.
- [13] We thank T. Bose for help with production of special MC samples.

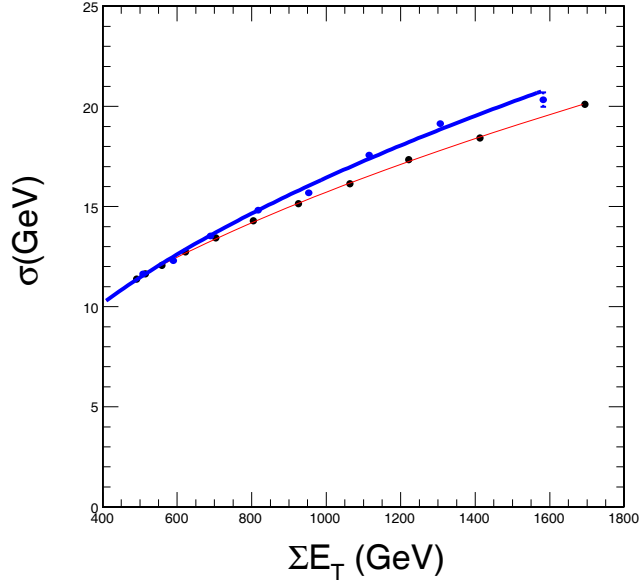


Figure 40: Distribution of $\sigma(\text{MET-perp})$ vs ΣE_T . The dark blue dots and (thick) fitted curve correspond to CMSSW_1.2.0 as studied in this note, whilst the black dots and the (thin) red fitted curve correspond to ORCA as presented in Fig. 11.17 of Ref. [1].

[14] CMS Note 2005/034, R. Harris, “*Jet Calibration from Dijet Balancing*”.

[15] H. Pi et al., CMS NOTE-2006/035, “*Measurement of Missing Transverse Energy With the CMS Detector at the LHC*”.